Global and regional aspects of tropical cyclone activity in the CMIP5 models

Suzana J. Camargo[1]
Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY

[1] *Corresponding author address*: Suzana J. Camargo, Lamont-Doherty Earth Observatory, Columbia University, PO Box 1000, Palisades, NY 10964-800
E-mail: suzana@ldeo.columbia.edu

Abstract

Tropical cyclone (TC) activity is analyzed in 8 Climate Model Intercomparison Project phase 5 (CMIP5) models. The global TC activity in the historical runs is compared with observations. Although all the models underestimate the global frequency of TCs, there is a wide range of in global TC frequency across the models. The models with the highest horizontal resolution have the highest level of global TC activity though resolution is not the only factor that determines TC activity. The models showed no consensus regarding the difference of TC activity in two warming scenarios (RCP4.5 and RCP8.5) and the historical simulation. We examined in detail North Atlantic and eastern North Pacific TC activity in a subset of models and found no robust changes across models in TC frequency. In two models, a westward moving track-type occurring in the eastern North Pacific and associated with El Niño had a higher frequency in the future scenarios. The future changes in various large-scale environmental fields associated with TC activity were also examined globally: genesis potential index, potential intensity, vertical wind shear and sea level pressure. The multi-model mean changes of these variables in the CMIP5 models are consistent with the changes obtained in the CMIP3 models.

1. Introduction

There is a huge interest in the potential change of tropical cyclone behavior with global warming due to the large impacts of tropical cyclones on coastal communities around the world. Three approaches to inferring tropical cyclone (TC) activity from climate models are: (i) to examine the statistics of TC-like storms generated by models, (ii) to analyze the large-scale variables associated with TC activity, or (iii) to perform statistical or dynamical downscaling. Each of these approaches has positive and negative aspects. In this paper, we will apply the first two approaches to data from the Climate Model Intercomparison Project phase 5 (CMIP5) dataset (Taylor et al. 2012).

The main objectives of this paper are to assess the ability of the CMIP5 climate models to simulation TC activity and to determine whether the models show robust global and regional TC activity responses in to warming. Our regional focus is on the basins affecting the North American climate, namely the North Atlantic and the eastern North Pacific.

Low-resolution climate models can generate TC-like structures (e.g. Manabe et al. 1970; Bengtsson et al. 1982; Vitart et al. 1997; Camargo et al. 2005). These model TCs have some characteristics similar to observed TCs, including temporal and spatial climatological distributions, but are much weaker and larger than observed storms due to the low-resolution. Even when the TC model mean frequency is not correctly simulated, these models capture interannual variability associated with El Niño – Southern Oscillation (ENSO), and have been used successfully to develop dynamical (Vitart and Stockdale 2001; Camargo and Barnston 2009) and statistical-dynamical (Wang et al. 2009) seasonal forecasts of TC activity. More recently, multi-year hurricane forecasts have been developed using these models (Smith et al. 2010; Vecchi et al. 2012).

In the last few years, many centers have started to use high-resolution global climate models with more realistic TC characteristics to simulate TC activity (e.g. Bengtsson et al. 2007a,b; Gualdi et al. 2008; Zhao et al. 2009) across a range of time-scales: including intra-seasonal (Vitart 2009; Vitart et al. 2010, Jiang et al. 2012), seasonal (Zhao, et al. 2009; Chen and Li, 2009), and longer time-scales, with considerable success (e.g. Oouchi et al. 2006; Chauvin et al. 2006, Sugi et al. 2009). In most cases, these simulations are forced with prescribed sea surface temperature (SST); only in a few cases are fully coupled models used (e.g. Gualdi et al. 2008). Even these high-resolution global models are not able to simulate the most intense storms, and downscaling methods (statistical, dynamical and statistical-dynamical) have been employed to obtain more precise information about projected TC characteristics, especially intensity (Knutson et al. 2008; Bender et al. 2010; Lavender and Walsh 2011, Zhao and Held 2010, Villarini and Vecchi, 2012a,b).

As low-resolution climate models are better able to simulate the large-scale environmental than individual storms, one attractive approach is to analyze large-scale variables known to be associated with TC activity, instead of model TCs directly. Gray (1979) first developed a genesis index based on 4 parameters associated with TC occurrence. Emanuel and Nolan (2004) improved the Gray index, and further refinements have been suggested by various authors (e.g. Emanuel 2010; Tippett et al. 2011). The simplicity of these indices is their main attraction, and they have been applied to analyze TC activity on various time-scales, including intraseasonal (Camargo et al. 2009), seasonal (Camargo et al 2007b), future climate change (Veechi and Soden 2007b), and

past climates (Korty et al. 2012a,b). Following the same principle used in the CMIP3 models, analysis of projected changes of TC activity were performed using large-scale environmental variables known to be associated with TC activity, such as potential intensity, vertical wind shear and humidity (Vechi and Soden a,b, here called VS07a and VS07b).

Given the scope of the CMIP5 experiment design, most modeling centers contributed output from fairly lower resolution models. Therefore, it is important to consider the large-scale environmental changes, in addition to model storms. We expect that the CMIP5 simulation of TC activity will not be as good as in high-resolution simulations, but we want to know how far are the CMIP5-class models are from the high-resolution ones. We are particularly interested in whether the CMIP5 models project robust changes in the global TC activity similar to those seen in the high-resolution projections.

Until now, projected changes in TC activity are robust only on a global scale, with an expected small reduction in global TC frequency and a small increase in TC intensity by the end of the 21st century (Knutson et al. 2010). We want to know if the CMIP5 models reproduce these projected changes in global frequency and intensity. We will also explore regional robust changes in TC activity across CMIP5 models focusing on the North Atlantic and eastern North Pacific.

A few of the results presented here also appear in the US National Oceanic and Atmospheric Administration (NOAA) Modeling Analysis and Prediction Program (MAPP) synthesis papers on North American climate in the CMIP5 models (Sheffield et al. 2012 and Maloney et al. 2012).

In section 2, we describe the models, data and methods used in this paper. Section 3 discusses the global TC activity in the CMIP5 models, with a detailed analysis of TCs in the eastern North Pacific and North Atlantic region presented in Section 4. The changes in large-scale environment from the end of the 20th century to the end of the 21st century are presented in Section 5. A discussion of the results of our analysis is given in Section 6.

2. Models, data and methods

a. CMIP5 models

The choice of models used in this analysis was based in their data availability in the CMIP5 data portals (Taylor et al. 2012). Tracking TC-like storms in the models using the Camargo and Zebiak (2002) algorithm requires 6-hourly environmental variables (namely: vorticity at 850hPa, temperature and winds in various pressure levels, and surface pressure), and this requirement was the main restriction in our model choices, since only a few models provided this output frequency at the time of our analysis.

Furthermore, the data had to be accessible for specific scenarios (more details in Taylor et al. 2012). Our analysis includes a historical simulation, and two future warming scenarios. The historical simulation is forced with observed atmospheric composition changes (natural and anthropogenic), as well as time evolving land cover. The historical simulations are available from the mid of the 19th century to the near present, but we restricted our analysis to the period 1950-2005. For the future scenarios, we chose two projection simulations forced with specified atmospheric concentrations, also called

"representative concentration pathways" (RCPs). The first one is a mid-range mitigation emissions scenario (RCP4.5), the second a high emissions scenario (RCP8.5). To include a model in our analysis, we required 6-hourly data for the historical run as well as either the RCP4.5 or RCP8.5 future scenarios. In Table 1 we list the models used in this study.

For the calculations based on monthly data (e.g. potential intensity, genesis potential index), we used all available ensemble members available, even if they were different than that used in tracking the cyclones. Unfortunately, one of the models had no monthly data available for our analysis. A list of the models and number of ensemble members used for 6-hourly and monthly data calculations are given in Table 2.

b. Data

The observed TCs data used in this paper are based on best-track datasets from the National Hurricane Center (North Atlantic and eastern North Pacific) and the Joint Typhoon Warning Center (western North Pacific, North Indian Ocean and southern hemisphere), and available online (Jarvinen et al. 1984; Neumann et al. 1999; Chu et al. 2002). NCEP/NOAA Reanalysis dataset was used in the calculation of the present climate genesis potential index (Kalnay et al. 1996; Kistler et al. 2001).

c. Methods

The Camargo-Zebiak detection and tracking algorithm was used to identify and track TC-like storms in the CMIP5 model output (Camargo and Zebiak, 2002). This algorithm has been used extensively in global (e.g. Camargo et al. 2005; Walsh et al. 2010; Kim et al. 2011) and regional climate models (Landman et al. 2005; Camargo et al. 2007c) and operationally in the IRI TC seasonal dynamical forecast (Camargo and Barnston 2009). This algorithm was slightly modified to use 850hPa wind speed instead of surface wind speed and 3 instead of 4 temperature levels (850hPa, 500hPa and 300hPa), due to their unavailability in the CMIP5 6-hourly data. The algorithm detects and tracks structures with local maximum vorticity (850hPa) and wind speed, minimum surface pressure, and a warm core (based on temperature and wind fields). To be defined as a model storm, the values of the vorticity, wind speed and local temperature anomaly have to be above model and basin-dependent thresholds and last at least 2 days. These thresholds are based on the statistics of historical runs. The same thresholds are used the future climate projection simulations. Once the storm passes these criteria, the track is extended in time backwards and forwards using a relaxed vorticity threshold.

The Genesis Potential Index (GPI) used here was developed in Emanuel and Nolan (2004) and was discussed in detail in Camargo et al. 2007b. The GPI has been extensively applied and analyzed (e.g. Camargo et al. 2007a; Nolan et al. 2007; VS07a; Camargo et al. 2009; Tippett et al. 2011; Menkes et al. 2012). The GPI is a measure of potential formation of TCs based on 4 environmental variables, namely, low-level vorticity, vertical wind shear, mid-level relative humidity and potential intensity. To facilitate comparison among the models, the GPI was calculated on a 2 deg x 2 deg grid for all models, with all model fields being first interpolated to this grid before the GPI was calculated.

The potential intensity (PI) is a theoretical limit for TC intensity (Emanuel, 1988). The procedure for the calculating PI calculation was first developed in Emanuel (1995)

and later modified to take into account dissipative heating (Bister and Emanuel 1998; 2002a,b). PI depends on sea surface temperature (SST), sea level pressure, and profiles of temperature and humidity. PI has been extensively used as a proxy of TC intensity in analysis of low-resolution climate models (e.g. VS07b; Camargo et al. 2012), as the local PI has a high correlation with actual TC intensities in various time scales (Emanuel 2000; Wing et al. 2007). Similarly to the case of the GPI, the PI was calculated on a 2 deg x 2 deg uniform grid for all models.

The cluster analysis used in our analysis was developed in Gaffney (2004) and is described in detail in Gaffney et al. (2007). The cluster technique constructs a mixture of quadratic regression models, which are used to fit the geographical shape of TC tracks. Finite mixture models are able to fit highly non-Gaussian probability density functions using few component probability distribution functions. The model is fit to the data by maximizing the likelihood of the parameters conditioned on the data. One important advantage of this method is that it easily accommodates tracks of different lengths. Each TC track is assigned to one of the clusters. The number of clusters to be used is not uniquely determined, the optimal choice depends on the log-likehood values (interpreted as goodness of fit), and within-cluster spread (distance of all tracks in the cluster to the mean regression track). As the number of clusters (K) increase, the log-likelihood values increase and the within-cluster spread decreases, but both curves show diminishing improvement in fit for K higher than a certain value, which leads to an optimal range of K choices. The final selection within this range is usually based on the knowledge of the system. This technique has been applied to observed TC tracks in various regions including the western North Pacific (Camargo et al. 2007d,e), the eastern North Pacific (Camargo et al. 2008), the Fiji islands (Chand and Walsh, 2009, 2010), the North Atlantic (Kossin et al. 2010), and more recently to the southern hemisphere (Ramsay et al. 2012).

3. Historical Global TC activity

Models tracks and first position locations in the 8 models for the period 1980-2005 are shown in Figs. 1 and 2, respectively. Only one ensemble member is shown for models with more than one ensemble member. The models present a wide range of levels of global TCs. Some models (CanESM2, FGOALS and NorESM1) have very few TC tracks overall with a tendency for TCs in high latitudes in the North Pacific and North Atlantic. These three models are relatively active in the South Pacific subtropics. In the deep tropics, CanESM2 and FGOALS are most active in the south Indian Ocean.

The other 5 models (GFDL, HadGEM2, MIROC5, MPI and MRI) present a much higher level of TC activity, but with significant differences among the models; all models have biases compared with observations. The GFDL model is active in the Pacific and Indian Oceans, with relatively fewer storms in the North Atlantic. Another interesting characteristic of the GFDL model is that TCs occur very close to the Equator in the Central and Western Pacific and Indian Ocean.

Similar to the GFDL model, there are very few North Atlantic storms in the HadGEM2 model, with most TC activity occurring in the southern hemisphere, western North Pacific and Bay of Bengal. The other 3 models (MIROC5, MPI, and MRI) are the most active models globally. However, much of the TC activity occurs in the subtropical region and in latitudes closer to the poles than in observations. One possible explanation for this behavior is that the tracking scheme is not distinguishing well between tropical

and extra-tropical storms, despite one of the criteria being that the storm have a warm core. However, this high level activity out of the tropics is not present in all models (e.g. HadGEM2 and GFDL), therefore this issue needs to be examined further in more detail.

The distribution of the global number of TCs per year in the period 1980-2005 in the models and in observations is given in Fig. 3. Typical of low-resolution climate models (Camargo et al. 2005; 2007a) all models have too few storms per year. There is a clear relationship between model horizontal resolution and TC activity level, with the models with highest horizontal resolution (MRI, see Table 1), being the most active and closest to observed values. However, resolution does not completely explain TC activity level. For instance, the GFDL model has a lower resolution than the MPI model (Table 1), but their global TC activity levels are quite similar.

A comparison with the TC activity in the CMIP3 models (Meehl et al. 2007) is discouraging, with little improvement in the CMIP5 model simulations of TC activity (Walsh et al. 2010; 2012). TC activity also was seen to be better in CMIP3 models with higher horizontal resolution. Simulated TC frequency increases with increasing resolution if all other factors are kept constant (Bengtsson et al., 1995; Murakami and Sugi, 2010). However, model resolution is not the only factor responsible the quality of the TC simulations. TC frequency and spatial distribution in climate models are sensitive to changes in model convection schemes (Vitart et al. 2001; Kim et al. 2011, Reed and Jablonowski, 2011a; Zhao et al. 2012). Walsh et al. (2012) pointed to dramatic changes in model TC frequency in two versions of the CMIP3 GFDL model with different dynamical cores and the same convection parameterizations. An extensive analysis of idealized simulations using different dynamical cores for the same model showed that the quality of the TC simulation was dependent on the interaction of the different model dynamical cores and moist convection parameterizations (Reed and Jablonowski, 2012). In summary, increasing model horizontal resolution is not sufficient to improve its simulation of TC frequency, as the model TC activity is sensitive to physical parameterizations and dynamical cores.

A few models (MIROC5, MPI, MRI) are very active in the South Atlantic basin, a region that in observations very few storms occur. Hurricane Catarina in 2004 was a very unusual event (Pezza and Simmonds, 2005; McTaggart-Cowan et al., 2006). The occurrence of South Atlantic hurricanes in climate models is not unusual (e.g. Gualdi et al. 2008), but the level of activity in the MPI and MRI models in that region is quite high.

The climatological mean GPI is shown in Fig. 4 for the 8 models and NCEP reanalysis for the period 1950 - 2005. As noted in previous studies (Camargo et al. 2007a, Tippett et al. 2011; Walsh et al. 2010, 2012), the GPI values in the models are much higher than in the reanalysis. These studies attributed the difference to the lower values of relative humidity in mid-levels in the reanalysis compared to the climate models, as there are known differences between the relative humidity in the ERA and NCEP reanalysis (Daoud et al. 2009) and biases in the mid-troposphere relative humidity in the NCEP reanalysis (Bony et al. 1997). We calculated the difference between the models and NCEP reanalysis for the annual zonal mean climatological relative humidity at 600hPa between 40S and 40S for the period 1950 – 2005, and all models, with exception of CanESM2, have larger values than the reanalysis. Note that CanESM2 is the model with the smallest values of GPI.

One aspect of interest is the relationship between GPI and the TC occurrence in models. Low-resolution climate models tend to have more realistic patterns of GPI than of TC occurrence and there is no good relationship between GPI and TC occurrence in the models (Camargo et al. 2007a; Walsh et al. 2012). The same is true here, while CanESM2 and NorESM have almost no TC occurrence in the tropics, and very different from observations, their GPI pattern is quite similar to the reanalysis. The MRI model, the model with the highest resolution, shows the most agreement between GPI and TC frequency, which is, in agreement with Walsh et al. 2012. Even in the South Atlantic this relationship holds, with the MRI GPI in that region being quite high, and the model producing many TCs.

The next issue we want to examine is the model projections of global TC frequency. Here we exclude from our analysis the 3 models (CanESM2, FGOALS, NorESM) with close to zero mean number of TCs (NTCs) globally (see Fig. 3). The distributions of global NTC per year in the historical runs and the 2 projection scenarios (RCP4.5 and RCP8.5) are shown in Fig. 5. There is little consistency among the models. While there is slight increase in NTC in the future for the GFDL and MPI models, there is a large increase in the MRI model, and a small decrease for the HadGEM2 and MIROC5 models. Knutson et al. (2010) analyzed the projections of global TC frequency in many high-resolution climate models, and the robust response among them was a small (but significant) decrease in the global frequency of TCs at the end of the 21st century. The lack of consistency among the models can be partly explained by the low-resolution and bias in NTC in the models analyzed here.

In the case of the MRI model, the model with highest resolution in this set of CMIP5 models, the horizontal resolution is still lower than those models discussed in Knutson et al. (2010). It is interesting that here the MRI model (Mizuta et al. 2012) projects an increase the frequency of TCs in the future, while previous results with the various versions of this atmospheric model projected a decrease in the global frequency of TCs when forced with fixed SSTs (Sugi et al. 2009; Sugi et al. 2012; Murakami et al. 2012a,b). The reasons for the differences in these results could have multiple sources: the lower resolution in the CMIP5 simulations, coupled ocean instead of fixed SSTs, coupling with chemical and carbon models, differences in the algorithm used for detection and tracking TCs (including thresholds definitions, e.g. Walsh et al 2007). Similarly, the MPI projections for an increase in global NTC though not as dramatic as in the case of MRI, are still in contrast with results with previous projections using a high-resolution version of the model (Bengtsson et al. 2007b).

4. TC activity in the North Atlantic and eastern North Pacific.

a. TC activity in the present climate

We would like to explore now in more detail the TC characteristics of these simulations in the North Atlantic (NATL) and the eastern North Pacific (ENP). Previous studies have shown that most low-resolution models have difficulty in simulating the mean NTC in those regions, even when they are able to simulate well the interannual variability (Bengtsson et al. 1995; Vitart et al. 1997; Camargo et al. 2005,2007b; Walsh et al. 2010),.

Projections of NATL TC activity have been the focus of many studies using high-resolution global climate models (Zhao et al. 2009), regional climate models (Knutson et al. 2008, Bender et al. 2010), statistical-dynamical downscaling (Emanuel et al., 2008). We would like to examine if the CMIP5 model regional projection of the TC activity Atlantic is robust and if it is in agreement with these studies. Recently, Villarini and Vecchi (2012b) used a statistical downscaling methodology to examine the 21st century projections of Atlantic storms using the CMIP5 models. They obtained an increase in the number of TCs in the 1st half of the 21st century, but obtained ambiguous results when the whole 21st century was considered. In the case of the ENP, there is no robust signal among models for that region (e.g., Emanuel et al. 2008).

Another aspect of TC activity we would like to investigate is the possibility of track changes in these 2 regions. Given that the TC landfall location is determined by its track, if there are significant changes in track types these could lead to significant changes in landfall frequency and location. Murakami and Wang (2010) examined possible changes in NATL storms using high-resolution simulations of the MRI model and found significant track changes with a decrease of TC occurrence in the western part of the basin and increase in the eastern part of the basin. They attributed these track changes to changes in genesis locations, not to changes in circulation. We will examine the Atlantic tracks in the low-resolution version of the CMIP5 MRI model and compare with their results.

The tracks of TCs in the ENP and NATL for the CMIP5 models and observations are given in Fig. 6. Similar to the results for global activity (Figs. 1 and 2), there is an enormous range in the level of TC activity among the models. Four models, CanESM2, FGOALS-g2, HadGEM2, and NorESM1, have very few storms in the two basins.. In the case of HadGEM2, this lack of TC activity is a regional rather than global feature, in contrast to the other 3 models which have a very low global NTC,. The HadGEM2 is much more active in the other regions (western North Pacific, North Indian Ocean and southern hemisphere) than in the ENP and NATL. Given the low level of activity of these 4 models in these 2 basins, we will focus our attention in the other 4 models.

The average NTC per month in the NATL and ENP in models and observations is shown in Fig. 7. In both basins all 4 models have too few TCs per year. Examination of the annual cycle reveals that the models produce to few TCs during the active season and, in contrast, too many TCs during the inactive season, when there are none or very few TCs in observations. This leads again to the question of whether some of the storms that the algorithm are detecting are extra-tropical or cold core storms, and this issue should be examined in more detail.

Given the large differences among these 4 models in NTCs (Fig. 7) and tracks (Fig. 6) in the Atlantic and ENP, it is useful to examine the internal variability of TC activity in the models that have more than ensemble member available. In particular, we would like to know if the TC activity in different ensemble members of the same model is more similar than among different models. Fig. 8 shows the tracks for 2 ensemble members of the MPI model, and the NTC distribution per year and the mean NTC per month for 3 MPI ensemble members in the historical runs. Fig. 9 shows similar plots for 5 MIROC5 ensemble members. By comparing Figs. 8 and 9, it is clear that though there is variability in number and track patterns among the ensemble members of the same model, these are much more similar among each other than to other models (compare

with Figs. 6 and 7). This result is in agreement with the assessment of ensemble member uncertainty in TC activity done by Reed and Jablonowski (2011b) , which concluded that the dominant differences were due to different model versions and resolutions, and not due to sensitivity to internal variability.

b. TC activity in future climates

We now compare the present TC activity with RCP4.5 and RCP8.5 projections of TC activity in the Atlantic and ENP. In Fig. 10 we show the distribution of TCs per year in the GFDL, MIROC5, MPI and MRI models in the present and future. In the NATL, there is a statistically significant reduction in the TC frequency in the MIROC and MPI models, while for the GFDL model the NTC distributions are very similar in the present and future. There is a significant NTC increase in the MRI models for the RCP8.5 scenario. In summary, there is no robust signal across models in changes of Atlantic NTC by the end of the 21st century. This is result is not completely unexpected given the differences in the NTC climatology in the Atlantic in the models analyzed here. Furthermore, using a statistical downscaling technique, no robust changes in Atlantic NTC was obtained (Villarini and Vecchi, 2012b).

In the ENP, the MRI model also has a significant increase in the RCP8.5 scenario, with the MPI having a small increase in the TC frequency in both future scenarios. The MIROC5 model has a significant decrease in NTC in both scenarios. As in the Atlantic, there is almost no robust change in the ENP NTC across the models.

c. Tracks in present and future climates

The climatological track distributions (frequency of TC tracks per grid point) for in the present climate for the 4 models in the Atlantic and ENP are shown in the left panels of Fig. 11. For the models with more than one ensemble member available, all ensembles were used in the construction of the track density (MPI, MIROC5). The GFDL model Atlantic tracks are restricted to a narrow region starting Africa and going northwestward in direction of North America. The MIROC5 Atlantic tracks are in the subtropics, with only a few tracks of storms that form in the deep tropics. The most realistic Atlantic tracks in this group of models occur in the MPI model, which include both tropical and subtropical storms. Similarly to the MIROC5 model, the tracks in the MRI model mainly occur in subtropical latitudes, with only a few cases in the deep tropics.

The tracks in the ENP (Figs. 6 and 11) are more similar among the models than in the case of the Atlantic, but there are clear differences as well. The tracks in the GFDL model are restricted to a smaller region, while the MRI model have tracks occurring nearer the Equator and more subtropical latitudes. The MPI and MIROC tracks tend to be mainly parallel to the coast. The difference of the track density in future and historical scenarios is shown in the central (RCP4.5) and right panels (RCP8.5) of Fig. 11. The largest differences occur in the MIROC5 model. However, these negative changes near the coast can be directly attributed to fewer storms in that model in the future. The 3 other models have noisier differences in the track density that we would like to examine in more detail using cluster analysis.

We will restrict our cluster analysis to the MPI and MRI models, as these models have more ensemble members available. Furthermore the tracks patterns of these models are not completely unrealistic, as we will show below. With cluster analysis, we can separate the TC tracks in groups with similar track types. As we have already applied the same cluster analysis for observed tracks in the Atlantic and ENP, we can compare the model track types with the observed ones as well as between models. We can also compare the cluster assignments in present and in future scenarios and identify track shifts in the future climates.

The cluster analysis was applied separately for the Atlantic and ENP storms. We used all the tracks in all scenarios and ensembles members to construct our cluster analysis. We re-ran the cluster analysis 100 times re-ordering the storms randomly to obtain the statistical significance of the cluster assignments (see the details in Camargo et al. 2007a, 2008, Kossin et al. 2010). Fig. 12 shows the result of the cluster analysis for the MPI (left panels) and MRI (right panels) Atlantic tracks. When this cluster analysis was applied to observed Atlantic tracks, the optimal number of cluster choice was K=4 (Kossin et al. 2010). Here we use the same choice for the number of clusters for the models. The 4 clusters in observations include a cluster of subtropical storms, one of Gulf of Mexico storms, and two types of deep tropics storms, one with formation more to the east of the basin and the other near the Caribbean islands (see Fig. 1 in Kossin et al. 2010).

The MPI model has 3 clusters in the Atlantic that are similar to the observed ones: a subtropical cluster, a Gulf of Mexico cluster and a tropical cluster. The remaining cluster, with subtropical storms located in the eastern part of the basin (Fig. 12(c)) has no observed counterpart. In the case of the MRI model, there are 2 cluster types in the Atlantic that are somewhat similar to observed ones, a subtropical cluster (Fig. 12(b)) and a Gulf of Mexico cluster (Fig. 12(h)), though this last cluster also includes the few storms forming in the MDR in that model. The main difference in the MRI model tracks and the observations is the existence of two additional subtropical clusters, with no correspondence in the observations, one in the central Atlantic (Fig. 12(d)), and one in the eastern Atlantic (Fig. 12(f)), which is similar to the MPI subtropical eastern cluster (Fig. 12(c)).

In table 3, we show the percentage of storms in each cluster for both models in the historical and the 2 future scenarios. In cases with statistically significant differences (using a t-test), the percentages are marked in bold. For the MPI model, there is a statistically significant increase in the percentage of storms in the 2 subtropical clusters (K=1 and K=2), and a statistically significant decrease in the percentage of storms in the tropical cluster (K=4). In the case of the MRI model, the shifts in the future are between storms in clusters (K=1) and (K=3), i.e., subtropical storms in the western and eastern part of the basin, with a decrease in the percentage of subtropical storms in the western region and an increase in the subtropical tracks in the eastern part of the basin (which do not exist in observed tracks).

The Atlantic tracks in the high-resolution version of the MRI model used in Murakami and Wang (2010) are much more similar to the observed Atlantic tracks. There are no central and western Atlantic subtropical tracks, as are seen here. Similar the zonal (east to west) shift noted here, Murakami and Wang (2010) also obtained a zonal shift,

with a decrease of tracks in the Gulf of Mexico and Caribbean and an increase in the MDR and western subtropical storms.

In the case of the ENP observed tracks, the cluster analysis led to an optimal choice of 3 clusters (see Fig. 5 of Camargo et al. 2008): two clusters that form very near the coast, one with tracks parallel and close to the coast, and the other forming slightly northward and moving more northwestward, and a third cluster that forms further from the coast and have a more westward track type. In Fig. 13, we show the result of applying the cluster analysis to the ENP tracks for both MPI and MRI model. The similarity of the ENP tracks types in the models with the observations and between the two models is better than in the Atlantic. Both models have one cluster that forms near the coast and moves almost parallel to the coast (Fig. 13(a) and (b)) and another that forms further from the coast and has a westward moving track (Fig. 13(c) and (d)). The clusters least similar to the observed ones are the third type (Fig. 13(e) and (f)), but even in this case the tracks do not seem as unrealistic as some of the model Atlantic clusters.

The percentage of tracks in each cluster in both models and all scenarios is given in Table 4. In this case, both models project an increase in the percentage of westward moving storms (K=2, Fig. 13(c) and (d), but differ as to which cluster will decrease. While the decrease for the MPI model occurs in the most populated cluster, with tracks parallel to the coast (K=1, Fig. 13(a)), for the MRI the decrease occurs in the other near coast cluster (K=3, Fig. 13(f)). The consistency of the increase in the westward moving track between the two models is intriguing. In observations, this type of track is associated with occurrence of El Niño events. It would be interesting to examine if there is an increase in frequency in El Niño events in these two models that could explain this track shift in the ENP.

5. Changes in the large-scale environment

We will examine now changes in the large-scale environment in the two future scenarios, compared with the historical simulations in all models. In Fig. 14 we show the difference in the climatological GPI in the future and present for 7 models and their model mean. The climatological GPI in the historical simulation for the same models was shown in Fig. 4. The model mean GPI difference is positive in most regions, with exception of the central south Pacific. An increase in the GPI in the future can be interpreted as an increase in the global TC frequency in these models in a future climate. However, in some cases model GPI is observed to increase even when the model NTCs decreases (Camargo et al. 2012b), therefore we need caution in interpreting this result. The model with the largest increase of GPI is the MIROC5 model, which has a decrease in the number of TCs globally in this scenario (Fig. 5) as well. In some models, the GPI difference patterns appear to be shifts in location, such as in the southern hemisphere and western North Pacific for the NorESM1 model. These shifts resemble GPI ENSO difference patterns (El Niño minus La Niña) discussed in Camargo et al. (2007b). Another interesting feature is the decrease of GPI in the MRI model in the south Atlantic, where there is an unrealistically high number of TCs in the historical run (Fig. 1(g)).

Vecchi and Soden (2007a) calculated the model mean difference of GPI in future and present climates for the CMIP3 models in the June – November period (Fig. 4(d) in VS07a)). In that case, there was also an increase in most of the northern hemisphere GPI in the ensemble mean with a maximum in the western North Pacific, with exception of

the ENP, where the GPI decreased. Here the model mean difference is positive in the whole northern hemisphere. It is interesting to note that there are only small increases in the GPI in the Atlantic MDR region and near Florida in both cases.

Another quantity of interest is the potential intensity (PI), which is the theoretical maximum of TCs. Fig. 15 shows the difference of the potential intensity in the RCP8.5 and historical scenarios for the individual models and model mean. The PI increases in most of the northern hemisphere in all models, with exception of a small region of the eastern part of the NATL and Pacific Oceans. In some models, the decrease in in PI in the NATL is restricted to a small region near Africa (MRI model), while in other models a larger region in the NATL has negative PI differences, including the Gulf of Mexico (MPI and CanESM2) and MDR region (CanESM2). In the southern hemisphere, there is an increase in PI in most regions. In all models there is a strong decrease in PI in southeast Pacific and Atlantic, regions where normally TCs do not occur. Vecchi and Soden (2007)a,b analysis of the PI of the CMIP3 models for the June – November season has a very similar pattern in the northern hemisphere to that shown here: an increase in PI in most of the northern hemisphere with one maximum near Hawaii and the other centered in the Equator near the data line, and a decrease in the PI in the NATL, which was attributed to changes in the remote SST (VS07b). The NATL negative region in the CMIP5 model mean PI differences is smaller and more restricted to the eastern Atlantic, than in the case of CMIP3. We repeated the calculation of the multi-model mean using 22 CMIP5 models, instead of only the 8 included in this study and the pattern obtained looks very similar to that in panel (h).

We also analyzed changes in the magnitude of the vertical wind shear in future and present climates (Fig. 15). All models have large regions of increased (decreased) vertical shear in the subtropical (tropical) latitudes of the southern hemisphere, which could lead to an equatorial shift of the TC activity in the southern hemisphere. In the northern hemisphere, most models show an increase in the vertical shear in the ENP and Caribbean region, extending in some cases into the Gulf of Mexico. In contrast, the eastern part of the Atlantic, western North Pacific and North Indian Ocean have a reduction of the vertical wind shear. Once more, the model mean pattern is extremely similar to that obtained in the CMIP3 models in VS07a for the northern hemisphere TC season. These changes in vertical shear are associated to the projected decrease in the Pacific Walker circulation (VS07a), while the near equatorial vertical shear weakening was related to the near-equatorial zonal overturning (VS07a,c; Vecchi et al., 2006) due to global thermodynamic constraints (Held and Soden, 2006)

The difference of sea level pressure (SLP) in the RCP8.5 and historical runs is shown in Fig. 17. The movement of TCs (i.e. their tracks) is largely determined by the ambient flow, or steering winds, with modifications due to the beta effect (Chan 2005). The steering winds are strongly related to the position and strength of the subtropical highs. Therefore, any future changes in the subtropical highs will be associated with shifts in TC tracks (Colbert and Soden 2012). The differences in SLP in the NATL shown in Fig. 17 are associated with a westward expansion of the subtropical high, which could potentially lead to more landfalls in the southeast region of the United States. Li et al. (2012) has noticed this extension of the NATL subtropical high in the CMIP5 models for the RCP4.5 scenario, and it was also present in the CMIP3 models (Li et al. 2011). In the western North Pacific, the increase in the SLP in the RCP8.5 projections could indicate a

southwestward shift of the subtropical high. Similarly, the increase of SLP in the southern hemisphere could be related to an Equatorial shift of the subtropical high in the south Pacific and south Indian Ocean.

6. Summary

An assessment of the TC activity in 8 CMIP5 models was presented. Although the typical model resolution increased since the previous CMIP3 assessment, model global TC frequency is still much lower than observed. Even in the model with the highest resolution, global TC occurrence is only half of the observed value. Furthermore, there are still deficiencies in the geographical patterns of the TC tracks and formation, with a large number of subtropical and South Atlantic storms in many models. There is no robust signal across the models in changes in global TC frequency for future scenarios. An analysis of the regional TC activity in the ENP and NATL did not detect any robust changes in TC frequency in those two regions either. We used cluster analysis to look for possible changes in tracks in these two regions, with consistent results obtained only in the ENP, with westward moving storms occurring more often than in the historical case in the 2 models.

Given the current state of the art of the CMIP5 models for TC activity, it is fundamental to continue to use a variety of downscaling methods (statistical and dynamical) to better infer future projections of TC frequency, intensity and tracks (e.g. Knutson et al. 2008; Villarini and Vecchi 2012a,b). High-resolution models forced with fixed SST from the CMIP5 models (e.g. Zhao et al. 2009) and statistical-dynamical downscaling results (e.g. Emanuel et al. 2008) should still give a better assessment of future track and frequency projections than using only low-resolution models.

The large-scale environmental variables projection changes associated with TC activity in the CMIP5 models are very consistent with the CMIP3 results. More detailed analysis of these changes is necessary, for instance the seasonality of PI, which had changes in the CMIP3 models (Sobel and Camargo, 2011). As basin wide PI trends can be larger than actual local PI trends (Kossin and Camargo, 2009), caution must be taken not to overestimate future trends when using basin wide large-scale variables.

Acknowledgments:

SJC would like to thank Haibo Liu and Naomi Naik for collecting, managing and serving the CMIP5 data at the Lamont-Doherty Earth Observatory of Columbia University and Michael Tippett for suggestions on the manuscript. SJC is supported by the National Oceanic and Atmospheric Administration (NOAA) Modeling Analysis and Prediction Program (MAPP) under grant NA11OAR4310093 and is part of the NOAA MAPP CMIP5 Task Force. We acknowledge the World Climate Research Programme's Working Group on Coupled Modelling, which is responsible for CMIP, and we thank the climate modeling groups (listed in Tables 1 and 2 of this paper) for producing and making available their model output. For CMIP the U.S. Department of Energy's Program for Climate Model Diagnosis and Intercomparison provides coordinating support and led development of software infrastructure in partnership with the Global Organization for Earth System Science Portals.

References

Bao, Q., and others: The Flexible Global Ocean-Atmosphere-Land System model Version: FGOALS-s2. Adv. Atmos. Sci., submitted.

Bender, M.A., T.R. Knutson, R. E. Tuleya, J.J. Sirutis, G.A. Vecchi, S.T. Garner, I.M. Held, 2010: Modeled impact of anthropogenic warming on the frequency of intense Atlantic hurricanes. *Science*, **327**, 454-458.

Bengtsson, L., H. Böttger, and M. Kanamitsu, 1982: Simulation of hurricane-type vortices in a general circulation model. *Tellus*, **34**, 440 – 457.

Bengtsson, L., M. Botzet, and M. Esch, 1995: Hurricane-type vortices in a general circulation model. *Tellus* **47A**: 175-196

Bengtsson, L., K. I. Hodges, and M. Esch, 2007a: Tropical cyclones in a T159 resolution global climate model: Comparison with observations and re-analysis. *Tellus*, **59A**, 396-416.

Bengtsson, L., K.I. Hodges, M. Esch, N. Keenlyside, L. Kornblueh, J.-J. Luo and T. Yamagata, 2007b: How many tropical cyclones change in a warmer climate? *Tellus,* **59A**, 539 – 561.

Bister, M., and K. A. Emanuel, 1998: Dissipative heating and hurricane intensity. Meteor. Atmos. Phys., 65, 223 - 240.

Bister, M., and K. A. Emanuel, 2002a: Low frequency variability of tropical cyclone potential intensity: 1. Interannual to interdecadal variability. J. Geophys. Res., 107, 4801.

Bister, M., and K. A. Emanuel, 2002b: Low frequency variability of tropical cyclone potential intensity: 2. Climatology for 1982-1995. J. Geophys. Res., 107, 4621.

Bony, S., Y. Sud, K.M. Lau, J. Susskind, and S. Saha, 1997: Comparison and satellite assessment of NASA/DAO and NCEP-NCAR reanalyses over tropical ocean: Atmospheric hydrology and radiation. *J. Climate,* **10**, 1441 – 1462.

Camargo, S.J., and A.G. Barnston, 2009: Experimental seasonal dynamical forecasts of tropical cyclone activity at IRI. *Wea. Forecasting*, **24**, 472 – 491.

Camargo, S. J., and S. E. Zebiak, 2002: Improving the detection and tracking of tropical cyclones in atmospheric general circulation models. *Wea. Forecasting*, **17**, 1152 –1162.

Camargo, S. J., A. G. Barnston, and S. E. Zebiak, 2005: A statistical assessment of tropical cyclone activity in atmospheric general circulation models. *Tellus,* **57A**, 589 – 604.

Camargo, S. J., K. A. Emanuel, and A. H. Sobel, 2007b: Use of a genesis potential index to diagnose ENSO effects on tropical cyclone genesis. *J. Climate*, **20**, 4819 - 4834.

Camargo, S. J., H. Li, L. Sun, 2007c: Feasibility study for downscaling seasonal tropical cyclone activity using the NCEP regional spectral model. *Int. J. Climate*, **27**, 311 - 325.

Camargo, S. J., A. H. Sobel, A. G. Barnston, and K. A. Emanuel, 2007a: Tropical cyclone genesis potential index in climate models. *Tellus*, **59A**, 428 – 443.

Camargo, S.J., M. Ting, and Y. Kushnir, 2012a: Influence of local and remote SST on North Atlantic potential intensity. *Clim. Dyn.*, in review.
Camargo, S.J., M.C. Wheeler, and A.H. Sobel, 2009: Diagnosis of the MJO modulation of tropical cyclogenesis using an empirical index. *J. Atmos. Sci*., **66**, 3061–3074,
Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007d: Cluster analysis of typhoon tracks: Part I: General Properties. *J. Climate*, **20**, 3635-3653.
Camargo, S. J., A. W. Robertson, S. J. Gaffney, P. Smyth, and M. Ghil, 2007e: Cluster analysis of typhoon tracks: Part II: Large-scale circulation and ENSO. *J. Climate*, **20**, 3654-3676.
Camargo, S. J., A. W. Robertson, A. G. Barnston, and M. Ghil, 2008: Clustering of eastern North Pacific tropical cyclone tracks: ENSO and MJO effects. *Geochem. Geophys. Geosys.*, **9**, Q06V05.
Camargo, S.J., M.K. Tippett, A.H. Sobel, G.A. Vecchi, M. Zhao, and I.M. Held, 2012b: Analysis of tropical cyclone genesis indices for climate change using the HIRAM model. Proc. 30th Conference on Hurricanes and Tropical Meteorology, Abstract 4B.1, Ponte Vedra Beach, FL, 16 – 20 April 2012; manuscript in preparation.
Chan, J. C. L., 2005: The physics of tropical cyclone motion. *Annu. Rev. Fluid Mech.*, **37**, 99-128.
Chand, S.S., and K.J.E. Walsh, 2009: Tropical cyclone activity in the Fiji region: Spatial patterns and relationship to large-scale circulation. *J. Climate*, **22**, 3877–3893.
Chand, S.S., and K.J.E. Walsh, 2010: The influence of the Madden-Julian Oscillation on tropical cyclone activity in the Fiji region. *J. Climate*, **23**, 868 -886.
Chauvin, F., J.-F. Royer and M. Déqué, 2006: Response of hurricane-type vortices to global warming as simulated by ARPEGE-Climat at high resolution. *Clim. Dyn.*, **27**, 377 - 399
Chen, J-H., and S.-J. Lin, 2011: The remarkable predictability of inter-annual variablity of Atlantic hurricanes during the past decade. *Geophys. Res. Lett.*, **38**, L11804.
Chu, J.-H., C.R. Sampson, A.S. Levine, and E. Fukada, 2002: The Joint Typhoon Warming Center tropical cyclone best-tracks, 1945-2000. NRL Technical Report, NRL/MR/7540-02-16.
Colbert, A.J., B. J. Soden, 2012: Climatological variations in North Atlantic tropical cyclone tracks. *J. Climate*, **25**, 657–673.
Daoud, A.B., E. Sauquet, M. Langand, C. Obled, and G. Bontron, 2009: Comparison of 850-hPa relative humidity between ERA-40 and NCEP/NCAR re-analyses: Detection of suspicious data in ERA-40. *Atmos. Sci. Lett.*, **10**, 43-47.
Donner, L.J., co-authors, 2011: The dynamical core, physical parameterizations, and basic simulation characteristics of the atmospheric component of the GFDL global coupled model CM3. *J. Climate,* **24**, 3484-3519.
Emanuel, K. A., 1988: The maximum intensity of hurricanes. *J. Atmos. Sci.*, **45**, 1143 - 1155.
Emanuel, K. A., 1995: Sensitivity of tropical cyclones to surface exchange coefficients and a revised steady-state model incorporating eye dynamics. *J. Atmos. Sci.*, **52**, 3969 - 3976.
Emanuel, K.A., 2000: A statistical analysis of hurricane intensity. Mon. Wea. Rev., 128, 1139 - 1152.

Emanuel, K., 2010: Tropical Cyclone Activity Downscaled from NOAA-CIRES Reanalysis, 1908-1958. *J. Adv. Model. Earth Syst*., **2**, doi:10.3894/JAMES.2010.2.1.
Emanuel, K. and D. S. Nolan, 2004: Tropical cyclone activity and the global climate system. *Proc. 26th AMS Conf. on Hurr. and Trop. Meteor.*, #10A.2, p. 240-241.
Emanuel, K., R. Sundararajan, and J. Williams, 2008: Hurricanes and global warming: Results from downscaling IPCC AR4 simulations. *Bull. Amer. Meteor. Soc.*, **89**, 347 – 367.
Gaffney, S.J., 2004: Probabilistic curve-aligned clustering and prediction with regression mixture models. Ph.D. dissertation, department of Computer Sciences, University of California, Irvine.
Gaffney, S.J, A.W. Robertson, P. Smyth, S.J. Camargo and M. Ghil, 2007: Probabilistic clustering of extratropical cyclones using regression mixture models. *Clim. Dyn.*, **29**, 434 - 440.
Gualdi, S., E. Scoccimarro, and A. Navarra, 2008: Changes in tropical cyclone activity due to global warming: Results from a high-resolution coupled general circulation model. *J. Climate*, **21**, 5204 – 5228.
Held, I.M., and B.J. Soden, 2006: Robust responses of the hydrological cycle to global warming. *J. Climate*, **19**, 5686-5699.
Jarvinen, B. R., C. J. Neumann, and M. A. S. Davis, 1984: A tropical cyclone data tape for the North Atlantic basin, 1886-1983: Contents, limitations, and uses. NOAA Technical Memorandum NWS NHC 22, Coral Gables, Florida, 21 pp.
Jiang, X., M. Zhao, and D.E. Waliser, 2012: Modulation of tropical cyclones over the eastern Pacific by intraseasonal variability simulated in an AGCM. J. Climate, in press.
Jones, C.D. and co-authors, 2011: The HadGEM2-ES implementation of CMIP5 centennial simulations. *Geosci. Model Dev.*, **4**, 543 – 570.
Kalnay, E., and co-authors, 1996: The NCEP/NCAR 40-year reanalysis project. *Bull. Amer. Meteor. Soc.*, **77**, 437 – 471.
Kim, D., A.H. Sobel, A.D. del Genio, Y. Chen, S.J. Camargo, M.-S. Yao, M. Kelley, and L. Nazarenko, 2012: The tropical subseasonal variability simulated in the NASA GISS general circulation model. J. Climate, 25, 4541 – 4659.
Kistler R., and co-authors, 2001: The NCEP-NCAR 50-year reanalysis: Monthly means CD-ROM and documentation. *Bull. Amer. Meteor. Soc.*, **82**, 247 – 267.
Knutson, T.R., J.L. McBride, J. Chan, K. Emanuel, G. Holland, C. Landsea, I. Held, J.P. Kossin, A.K. Srivastava, and M. Sugi, 2010: Tropical cyclones and climate change. *Nature Geoscience*, **3**, 157-163.
Knutson, T. R., J. J. Sirutis, S. T. Garner, G. A. Vecchi, and I. M. Held, 2008: Simulated reduction in Atlantic hurricane frequency under twenty-first-century warming conditions. *Nature Geoscience*, **1**, 359-364.
Korty, R.L., S.J. Camargo, and J. Galewsky, 2012a. Tropical cyclone genesis factors in simulations of the Last Glacial Maximum. *J. Climate*, **25**, 4348-4365.
Korty, R.L., S.J. Camargo, and J. Galewsky, 2012b. Variations in tropical cyclone genesis factors in simulations of the Holocene Epoch. *J. Climate*, in press.
Kossin, J.P., and S. J. Camargo, 2009: Hurricane track variability and secular potential intensity trends. *Climatic Change*, **9**, 329 - 337.
Kossin, J.P., S. J. Camargo, and M. Sitkowski, 2010. Climate modulation of North Atlantic hurricane tracks. *J. Climate*, **23**, 3057-3076.

Landman, W. A., A. Seth, and S. J. Camargo, 2005: The effect of regional climate model domain choice on the simulation of tropical cyclone-like vortices in the Southwestern Indian Ocean. *J. Climate*, **18**, 1253-1274.
LaRow, T. E., Y.-K. Lim, D. W. Shin, E. P. Chassignet, and S. Cocke, 2008: Atlantic basin seasonal hurricane simulations, *J. Climate*, **21**, 3191-3206.
Lavender, S. L., and K. J. E. Walsh, 2011: Dynamically downscaled simulations of Australian region tropical cyclones in current and future climates, *Geophys. Res. Lett.*, 38, L10705.
Li, L., W. Li, and Y. Deng, 2012: Summer rainfall variability over the southeastern United States in the 21st century as assessed by the CMIP5 models. J. Geophys. Res., submitted.
Li, W., L. Li, R. Fu, Y. Deng, and H. Wang, 2011: Changes in the North Atlantic subtropical high and its role in the intensity of summer rainfall variability in the southeastern United States. *J. Climate*, **24**, 1499 – 1506.
Maloney, E.D., S.J. Camargo, E. Chang, B. Colle, R. Fu, K.L. Geil, Q. Hu, X. Jiang, N. Johnson, K.B. Karnauskas, J. Kinter, B. Kirtman, S. Kumar, B. Langenbrunner, K. Lombardo, L. Long, A. Mariotti, J.E. Meyerson, K. Mo, J.D. Neelin, Z. Pao, R. Seager, Y. Serra, A. Seth, J. Sheffield, J. Thibeault, S.-P. Xie, C. Wang, B. Wyman, and M. Zhao, 2012: North American climate in CMIP5 experiments: Part III: Assessment of 21st century projections. *J. Climate*, submitted.
Manabe, S., J.L. Holloway, and H.M. Stone, 1970: Tropical circulation in a time-integration of a global model atmosphere. *J. Atmos. Sci.*, **27**, 580 – 613.
McTaggart-Cowan, R., L.F. Bosart, C.A. Davis, E.H. Atallah, J.R. Gyakum, and K.A. Emanuel, 2006: Analysis of Hurricane Catarina (2004). *Mon. Wea. Rev.*, **134**, 3029 – 3053.
Menkes, C.E., M. Lengaigne, P. Marchesiello, N.C. Jourdain, E.M. Vincent, J. Lefevre, F. Chauvin, J.-F. Royer, 2011: Comparison of tropical cyclogenesis indices on seasonal to interannual timescales. *Clim. Dyn.*, **38**, 301 – 302.
Mizuta, R., and co-authors, 2012: Climate simulations using the improved MRI-AGCM with 20km grid.
Murakami, H., and M. Sugi, 2010: Effect of model resolution on tropical cyclone climate projections. *SOLA*, **6**, 73-76, doi: 10.2151/sola.2010-019.
Murakami, H., R. Mizuta, and E. Shindo, 2012a: Future changes in tropical cyclone activity project by multi-physics and multi-SST ensemble experiments using 60-km-mesh MRI-AGCM. Clim. Dyn., early online, doi: 10.1007/s00382-011-1223-x.
Murakami, H., and co-authors, 2012b: Future changes in tropical cyclone activity projected by the new high-resolution MRI-AGCM. *J. Climate*, **25**, 3237–3260.
Neumann, C.J., B.R. Jarvinen, C.J. McAdie, and G.R. Hammer, 1999: Tropical Cyclones of the North Atlantic Ocean, 1871-1998. Prepared by the National Climatic Data Center, Asheville, NC and the Tropical Prediction Center/National Hurricane Center, Miami, FL, 206 pp.
Nolan, D. S., E. D. Rappin, and K. A. Emanuel, 2007: Tropical cyclogenesis sensitivity to environmental parameters in radiative-convective equilibrium. *Q. J. R. Meteorol. Soc.*, **133**, 2085 - 2107.
Oouchi, K., J.Yoshimura, H. Yoshimura, R. Mizuta, S. Kusunoki, and A. Noda, 2006: Tropical cyclone climatology in a global-warming climate as simulated in a 20km-mesh

global atmospheric model: frequency and wind intensity analysis. *J. Meteorol. Soc. Japan*, **84**, 259-276.
Pezza, A. B., and I. Simmonds, 2005: The first South Atlantic hurricane: unprecedented blocking, low shear, and climate change. *Geophys. Res. Lett.,* **32**, L15712.
Ramsay, H.A., S.J. Camargo, and D. Kim, 2012: Cluster analysis of tropical cyclone tracks in the southern hemisphere. *Clim. Dyn.*, early online, doi: 10.1175/JCLI-D-12-00031.1.
Reed, K.A., and C. Jablownowski, 2011a: Impact of physical parametrization on idealized tropical cyclones in the Community Atmosphere Model. *Geophys. Res. Lett*., **38**, L048045.
Reed, K.A., and C. Jablownowski, 2011b: Assessing the Uncertainty of Tropical Cyclone Simulations in NCAR's Community Atmosphere Model. *J. Adv. Model. Earth Syst.*, **3**, M08002, 16, doi:10.1029/2011MS000076.
Reed, K.A., and C. Jablownowski, 2012: Idealized tropical cyclone simulations of intermediate complexity: A test case for AGCMs. *J. Adv. Model Earth. Sys.*, **4**, doi: 10.1029/2011MS000099.
Sheffield, J., S.J. Camargo, R. Fu, Q. Hu, X. Jiang, N. Johnson, K.B. Karnauskas, J. Kinter, S. Kumar, B. Langebrunner, E. Maloney, A. Mariotti, J. E. Meyerson, J.D. Neelin, Z. Pan, A. Ruiz-Barradas, R. Seager, Y.L. Serra, D.-Z. Sun, C. Wang, S.-P. Xie, J.-Y. Yu, T. Zhang, and M. Zhao, 2012: North-American climate in CMIP5 experiments: Part II: Evaluation of 20th century intra-seasonal to decadal variability. *J. Climate*, submitted.
Smith, D., R. Eade, N.J. Dunstone, D. Fereday, J.M. Murphy, H. Pohlmann, and A. A. Scaife, 2010: Skilfull multi-year predictions of Atlantic hurricane frequency. Nature Geoscience, 3, 846 – 849.
Sobel, A.H., and S.J. Camargo, 2011: Projected future changes in tropical summer climate. *J. Climate*, **24**, 473-487.
Sugi, M., H. Murakami, and J. Yoshimura, 2009: A reduction in global tropical cyclone frequency due to global warming. *SOLA*, **5**, 164 –167.
Sugi, M., H. Murakami, and J. Yoshimura, 2012: On the Mechanism of tropical cyclone frequency changes due to global warming. *J. Meteor. Soc. Japan*, **90A**, 397 – 408.
Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2012: An overview of CMIP5 and the experiment design. *Bull. Am. Meteorol. Soc.*, **93**, 485 – 498.
Tippett, M.K., S.J. Camargo, and A.H. Sobel, 2011: A Poisson regression index for tropical cyclone genesis and the role of large-scale vorticity in genesis. J. Climate, 24, 2335 - 2357.
Vecchi, G.A. and B.J. Soden 2007a: Increased tropical Atlantic wind shear in model projections of global warming. *Geophys. Res. Lett.*, **34**, L08702.
Vecchi, G.A., and B.J. Soden, 2007b: Effect of remote sea surface temperature change on tropical cyclone potential intensity. *Nature*, **450**, 1066-1070.
Vecchi, G.A., and B.J. Soden 2007c: Global warming and the weakening of the tropical circulation. *J. Climate*, **20**, 4316-4340.
Vecchi, G.A., B.J. Soden, A.T. Wittenberg, I.M. Held, A. Leetmaa, and M.J. Harrison, 2006: Weakening of tropical Pacific atmospheric circulation due to anthropogenic forcing. *Nature*, **441**, 73-76.

Vecchi, G.A., and co-authors, 2012: Multi-year predictions of North Atlantic hurricane frequency: Promise and limitations. J. Climate, submitted.
Villarini, G., and G.A. Vecchi, 2012a: Projected increases in North Atlantic tropical cyclone intensity from CMIP5 models. J. Climate, submitted.
Villarini, G. and G.A. Vecchi, 2012b: Twenty-first-century projections of North Atlantic tropical storms from CMIP5 models. Nature Clim. Change 2, 604 – 607.
Vitart, F., 2009: Impact of the Madden Julian Oscillation on tropical storms and risk of landfall in the ECMWF forecast system. *Geophys. Res. Lett.*, **36**, L15802.
Vitart F. D., and T. N. Stockdale, 2001: Seasonal forecasting of tropical storms using coupled GCM integrations. *Mon. Wea. Rev.*, **129**, 2521–2537.
Vitart, F., J. L. Anderson, and W. F. Stern, 1997: Simulation of interannual variability of tropical storm frequency in an ensemble of GCM integrations. *J. Climate*, **10**, 745 –760.
Vitart, F., J.L. Anderson, J. Sirutis, and R.E. Tuleya, 2001: Sensitivity of tropical storms simulated by a general circulation model to changes in cumulus parametrization. *Q. J. Royal Meteor. Soc.*, **127**, 25 - 51.
Vitart, F., A. Leroy, and M.C. Wheeler, 2010: A comparison of dynamical and statistical predictions of weekly tropical cyclone activity in the southern hemisphere. *Mon. Wea. Rev.*, **138**, 3671 – 3682.
Walsh, K. J., M. Fiorino, C. W. Landsea, and K. L. McInnes, 2007: Objectively determined resolution-dependent threshold criteria for the detection of tropical cyclones in climate model and reanalyses. *J. Climate*, **20**, 2307 – 2314.
Walsh, K., S. Lavender, H. Murakami, E. Scoccimarro, L.-P. Caron, and M. Ghantous, 2010: The tropical cyclone climate model intercomparison, in Hurricanes and Climate Change: Vol. 2, J. B. Elsner, R.E. Hodges, J.C. Malmstadt, and K.N. Scheitlin, editors, Springer.
Walsh, K., S. Lavender, E. Scoccimarro, and H. Murakami, 2012: Resolution dependence of tropical cyclone formation in CMIP3 and finer resolution models. Clim. Dyn., early online, doi: 10.1007/s00382-012-01298-z.
Wang, H., J.-K. Schemm, A. Kumar, W. Wang, L. Long, M. Chelliah, G.D. Bell, and P. Peng, 2009: A statistical forecast model for Atlantic seasonal hurricane activity based on the NCEP dynamical seasonal forecast. *J. Climate*, **22**, 4481 – 4500.
Watanabe, M. and co-authors, 2010: Improved climate simulation by MIROC5: Mean states, variability, and climate sensitivity. *J. Climate*, **23**, 6312-6335.
Wing, A. A., A. H. Sobel, and S. J. Camargo, 2007: The relationship between the potential and actual intensities of tropical cyclones on interannual time scales. *Geophys. Res. Lett.*, **34**, L08810.
Yukimoto, S. and co-authors, 2012: A new global climate model of the Meteorological Research Institute: MRI-CGCM3 – Model description and basic performance. *J. Meteor. Soc. Japan*, **90A**, 23-64, doi: 10.2151/jmsj.2012-A02.
Zanchettin, D., A. Rubino, D. Matei, O. Bothe, and J. H. Jungclaus, 2012: Multidecadal-to-centennial SST variability in the MPI-ESM simulation ensemble for the last millennium. *Clim. Dyn.*, **39**, 419-444 doi:10.1007/s00382-012-1361-9.
Zhang, Z. S., Nisancioglu, K., Bentsen, M., Tjiputra, J., Bethke, I., Yan, Q., Risebrobakken, B., Andersson, C., and Jansen, E., 2012: Pre-industrial and mid-Pliocene simulations with NorESM-L. *Geosci. Model Dev.*, **5,** 523-533, doi:10.5194/gmd-5-523-2012.

Zhao, M., and I. M. Held, 2010: An analysis of the effect of global warming on the intensity of Atlantic hurricanes using a GCM with statistical refinement. *J. Climate*, **23**, 6382-6393
Zhao, M., I.M. Held, and S.-J. Lin, 2012: Some counter-intuitive dependencies of tropical cyclone frequency on parameters in a GCM. *J. Atmos. Sci.,* doi: 10.1175/JAS-D-11-0238.1, in press.
Zhao, M., I.M. Held, S.-J. Lin, and G.A. Vecchi, 2009: Simulations of global hurricane climatology, interannual variability, and response to global warming using a 50km resolution GCM. *J. Climate*, **22**, 6653-6678.

TABLE 1. CMIP5 models used to track TC-like structures using 6-hourly data: official name, shortened name used in this paper in parentheses, modeling center, approximate horizontal resolution (in degrees) (Taylor et al. 2012).

Model Name	Modeling Center	Resolution	Reference or link:
CanESM2	Canadian Centre for Climate Modeling and Analysis	2.8 x 2.9	www.ec.gc.ca
FGOALS-g2 (FGOALS)	LASG, Institute of Atmospheric Physics, Chinese Academy of Sciences and CESS, Tsinghua University	2.8 x 3.0	Bao et al. (2012)
GFDL-ESM2M (GFDL)	NOAA Geophysical Fluid Dynamics Laboratory	2.5 x 2.0	Donner et al. (2011)
HadGEM2-ES (HadGEM2)	Met Office Hadley Center	1.9 x 1.2	Jones et al. (2011)
MIROC5 (MIROC)	Japan Agency for Marine-Earth Science and Technology, Atmosphere and Ocean Research Inst. (U. of Tokyo) and National Inst. for Environmental Studies	1.4 x 1.4	Watanabe et al. (2010)
MPI-ESM-LR (MPI)	Max-Planck Institute for Meteorology	1.9 x 1.9	Zanchettin et al. (2012)
MRI-CGCM3 (MRI)	Meteorological Research Institute	1.1 x 1.2	Yukimoto et al. (2011)
NorESM1-M (NorESM1)	Norwegian Climate Centre	2.5 x1.9	Zhang et al. (2012)

TABLE 2. CMIP5 models and number of ensemble members used to calculate six-hourly and monthly mean environmental variables

	Storms tracking – 6 hourly data			Environmental Fields – Monthly data	
Model Name	Historical	RCP4.5	RCP8.5	Historical	RCP8.5
CanESM2	1	1	1	5	5
FGOALS-g2	2	0	1	0	0
GFDL-ESM2M	1	1	1	1	1
HadGEM2-ES	1	1	1	4	4
MIROC5	1	3	2	4	3
MPI-ESM-LR	3	3	3	3	3
MRI-CGCM3	5	1	1	4	1
NorESM1-M	1	1	1	3	1

TABLE 3. The percentage of storms per cluster in the North Atlantic for the MRI and MPI models are given below. The percentages were calculating by giving each storm a cluster assignment, and repeating the procedure 100 times, randomizing the order in which the storms were entered in the cluster analysis. Percentages marked in bold show statistically significant differences for that cluster between the percentages in the historical and the correspondent scenario (RCP4.5 and RCP8.5) using a t-test

Model	Scenario	K=1	K=2	K=3	K=4
MPI	Historical	**0.33**	**0.27**	0.17	**0.22**
	RCP4.5	**0.35**	**0.31**	0.20	**0.13**
	RCP8.5	**0.37**	**0.30**	0.18	**0.13**
MRI	Historical	**0.32**	0.32	**0.23**	0.13
	RCP4.5	**0.33**	0.31	**0.25**	0.10
	RCP8.5	**0.29**	0.30	**0.27**	0.14

TABLE 4. The percentage of storms per cluster in the eastern North Pacific for the MRI and MPI models are shown below. The percentages were calculating by giving each storm a cluster assignment, and repeating the procedure 100 times, randomizing the order in which the storms were entered in the cluster analysis. All ensemble members available were included in the cluster analysis. Percentages marked in bold show statistically significant differences for that cluster between the percentages in the historical and the correspondent scenario (RCP4.5 and RCP8.5) using a t-test.

Model	Scenario	K=1	K=2	K=3
MPI	Historical	**0.59**	**0.18**	0.23
	RCP4.5	**0.54**	**0.28**	0.18
	RCP8.5	**0.50**	**0.34**	0.15
MRI	Historical	0.36	**0.31**	**0.33**
	RCP4.5	0.40	**0.32**	**0.29**
	RCP8.5	0.34	**0.37**	**0.29**

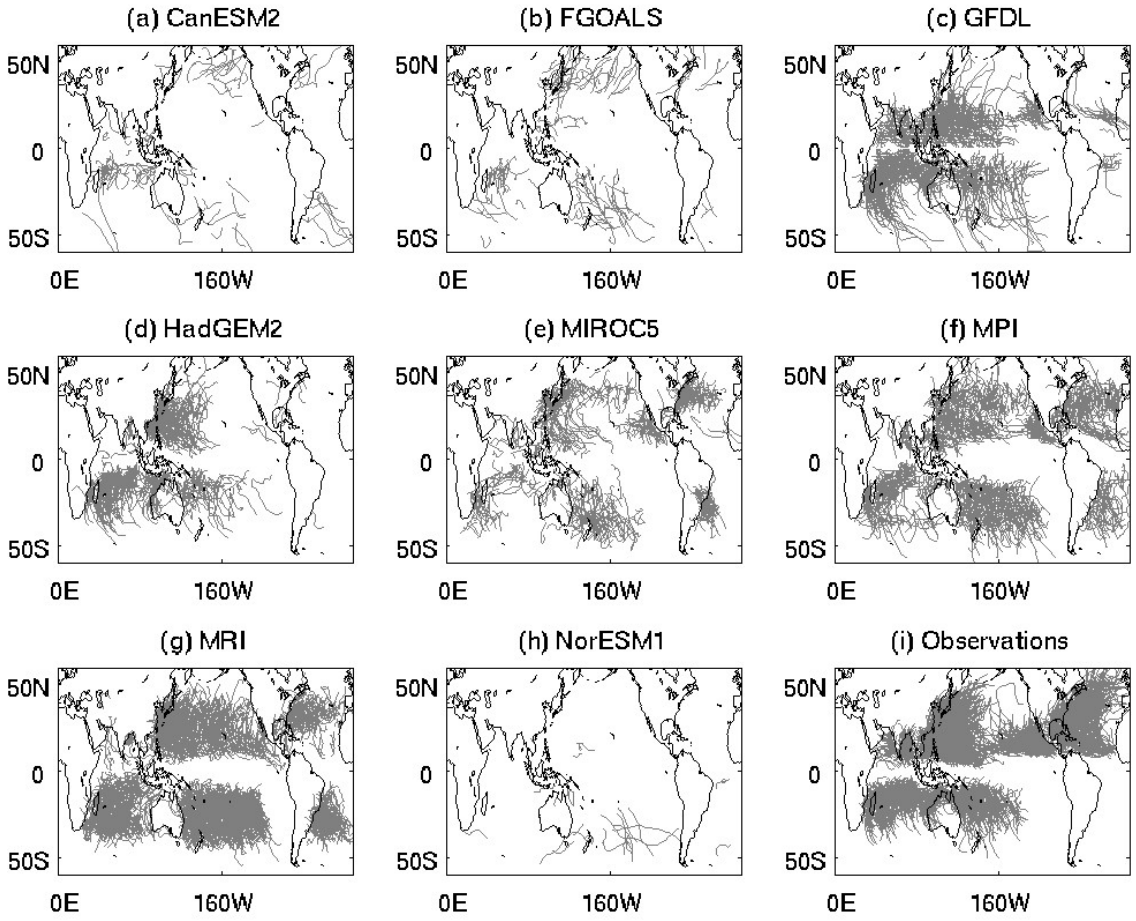

FIG. 1. TC tracks in 8 CMIP5 models (historical runs) and in observations for the period 1980-2005. Only one ensemble member is shown for each model.

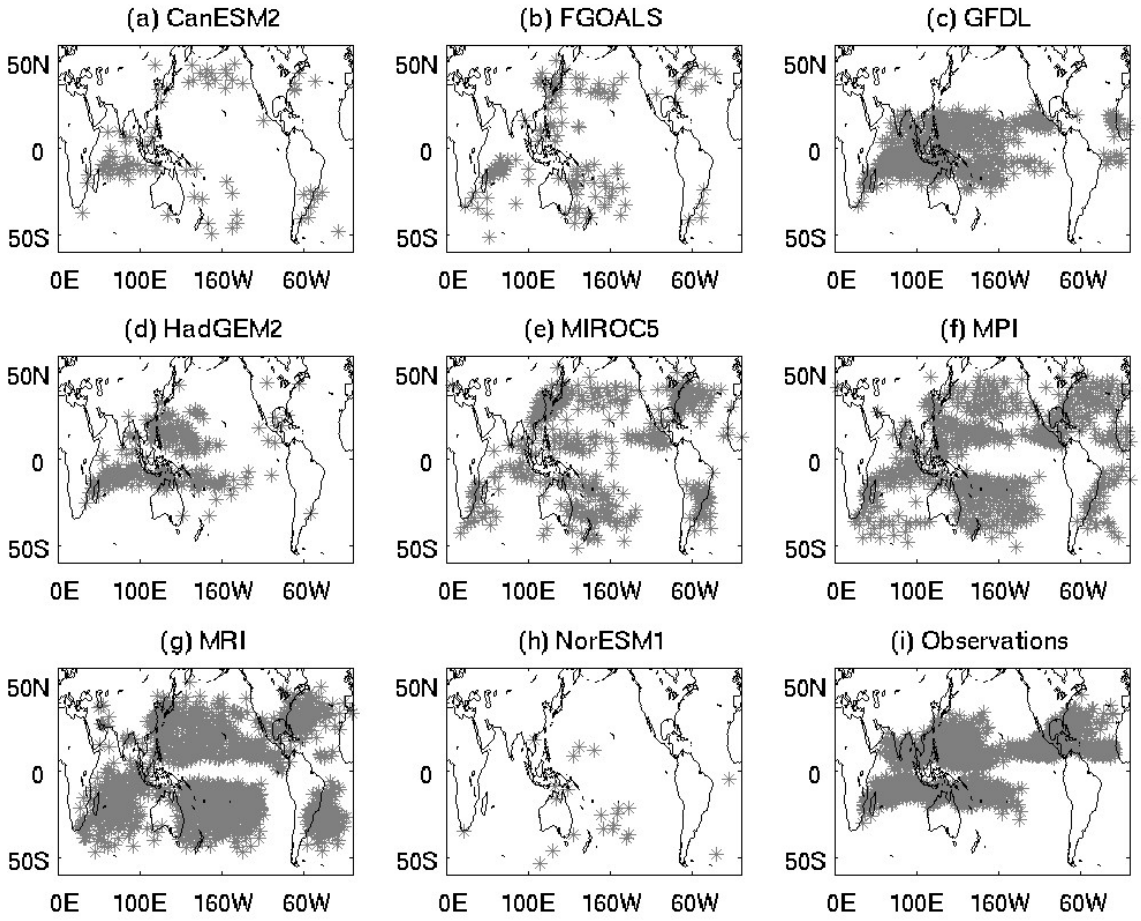

FIG. 2. TC first position in the tracks of 8 CMIP5 models (historical runs) and observations for the period 1980-2005 (shown in Fig. 1). Only one ensemble member is shown for each model.

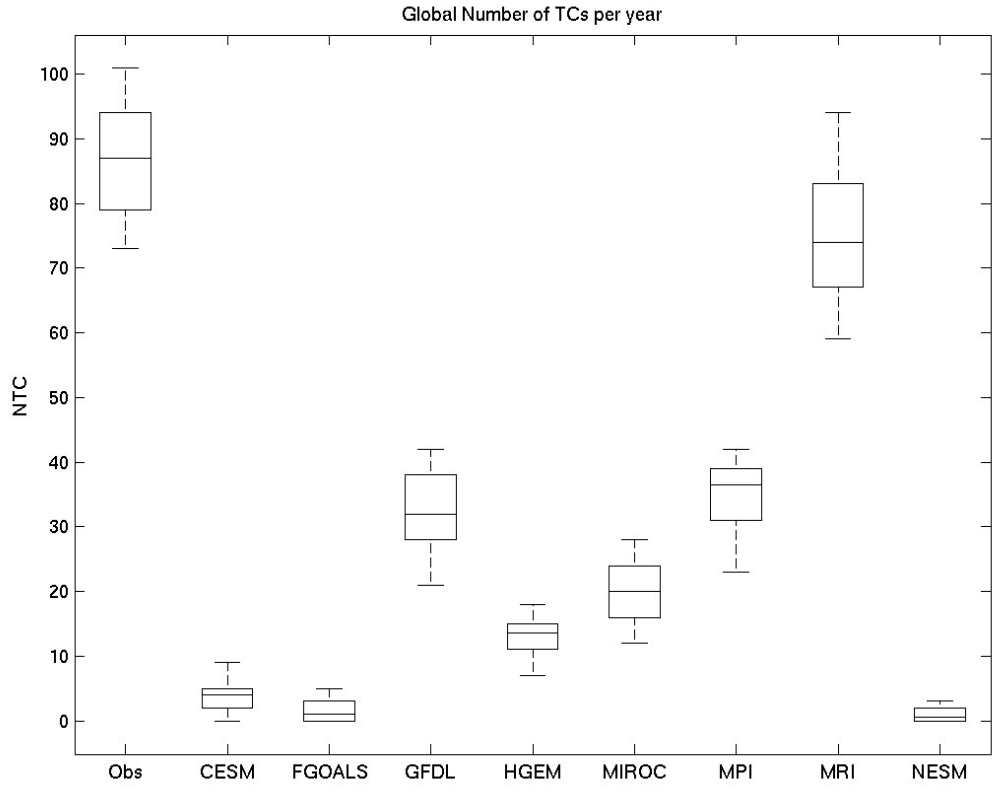

FIG. 3: Distribution of the global number of TCs per year in each of the models for the historical runs and in observations in the period 1980-2005: CanESM2 (CESM), HadGEM2 (HGEM), NorESM1 (NESM). When more than one ensemble member is available, all ensemble members are used in the distribution (FGOALS, MPI, MRI). The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

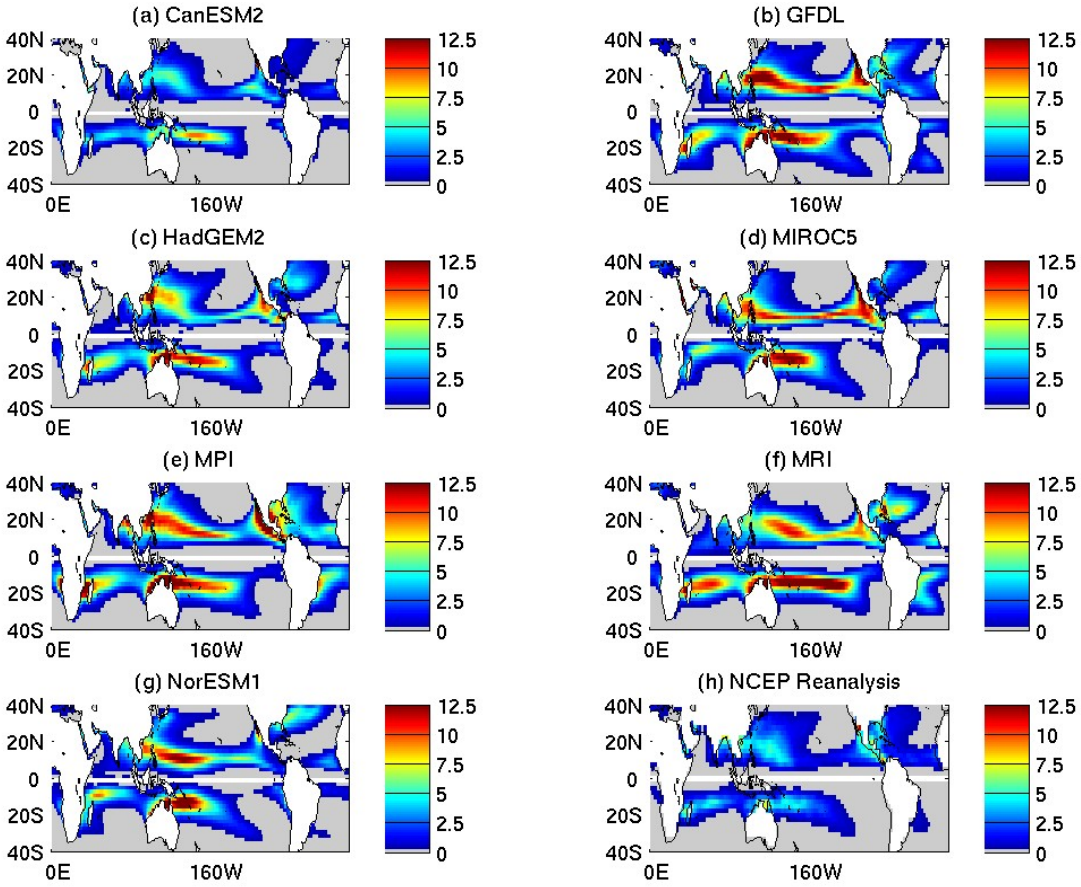

FIG. 4: Seasonal climatological Genesis Potential Index (GPI) in the models and the NCEP reanalysis for the period 1950-2005. In all panels, the northern (southern) hemisphere shows the GPI seasonal mean for August to October, or ASO (January to March, or JFM) season.

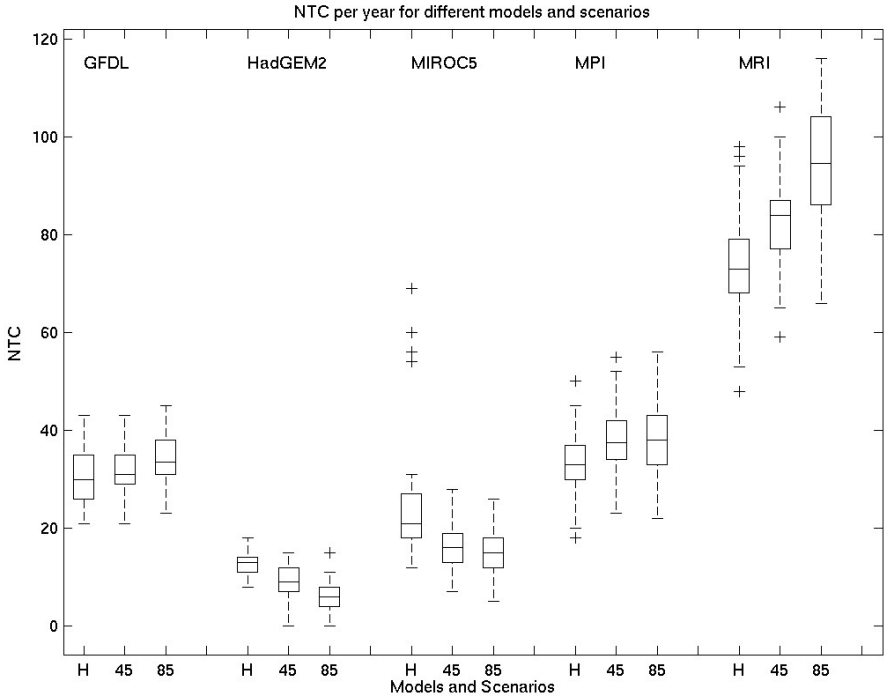

FIG. 5: Global number of TCs (NTC) per year in models for the historical (H) in the period 1951-2000 and in the future scenarios RCP4.5 (45) and RCP8.5 (85) in the period 2051-2100. The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

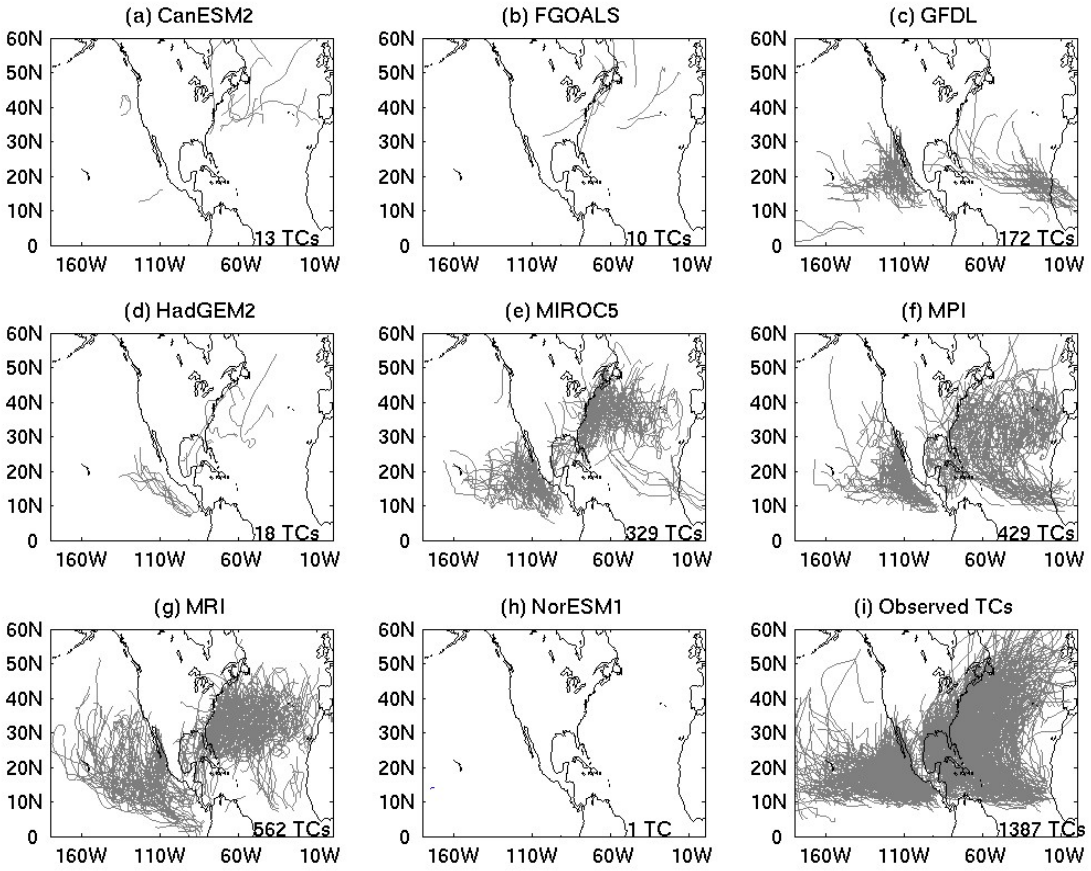

FIG. 6: North Atlantic and eastern North Pacific TC tracks in 8 CMIP5 models (historical runs) and in observations for the period 1980-2005. Only one ensemble member is shown for each model.

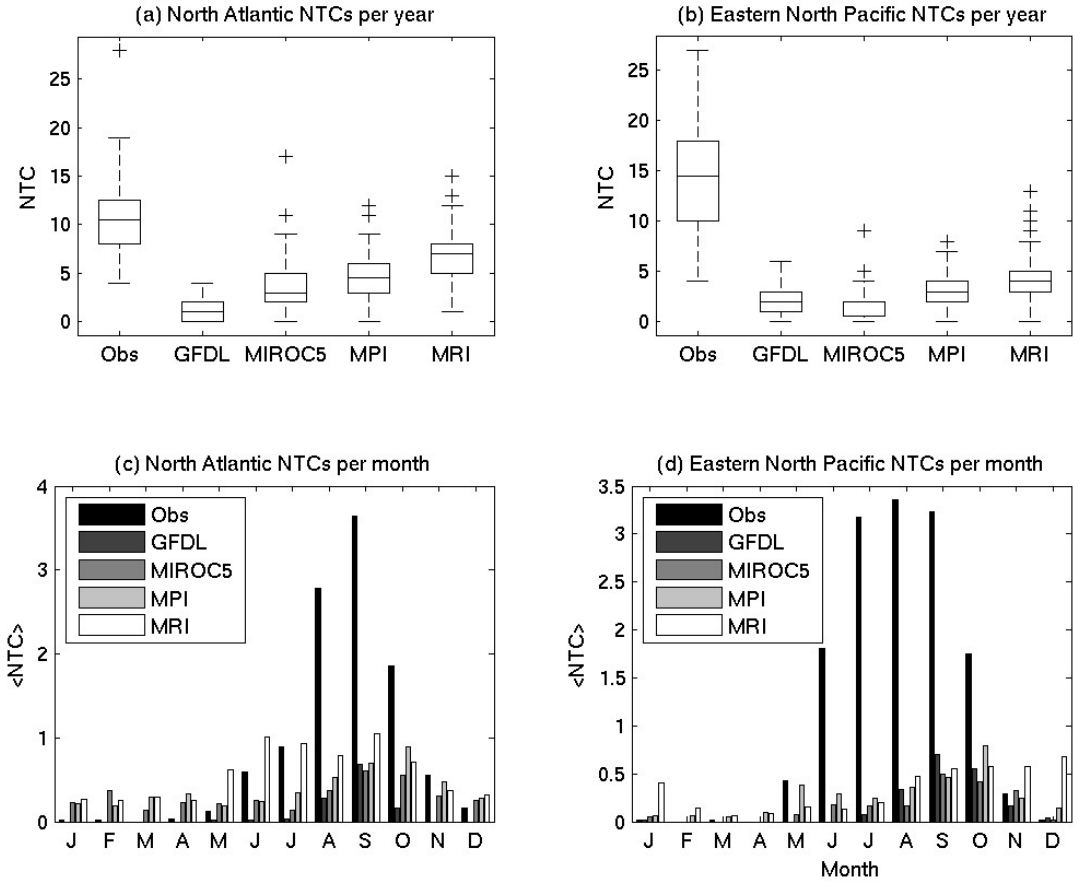

FIG. 7: Distribution of the number of TCs (NTC) in North Atlantic and eastern North Pacific in the period 1950-2005 for models and observation per year (top panels) and mean NTC per month (bottom panels). The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

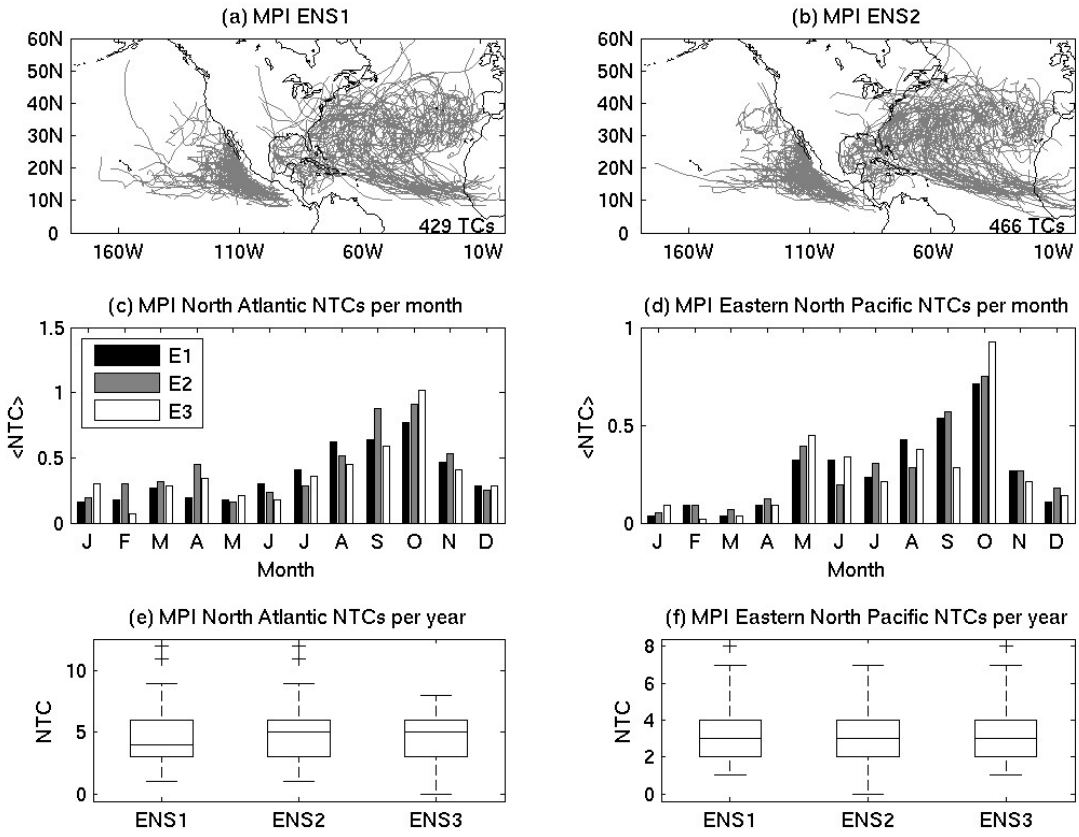

FIG. 8: Tracks of the MPI model TCs for ensemble members (ENS) 1 (a) and 2 (b). Mean MPI NTC per month in North Atlantic (c) and eastern North Pacific (d) for 3 ensemble members (E1, E2, and E3) in the period 1950-2005. Distributions of MPI NTC per year for 3 ensemble members in the North Atlantic (e) and eastern North Pacific (f) in the period 1950-2005. The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

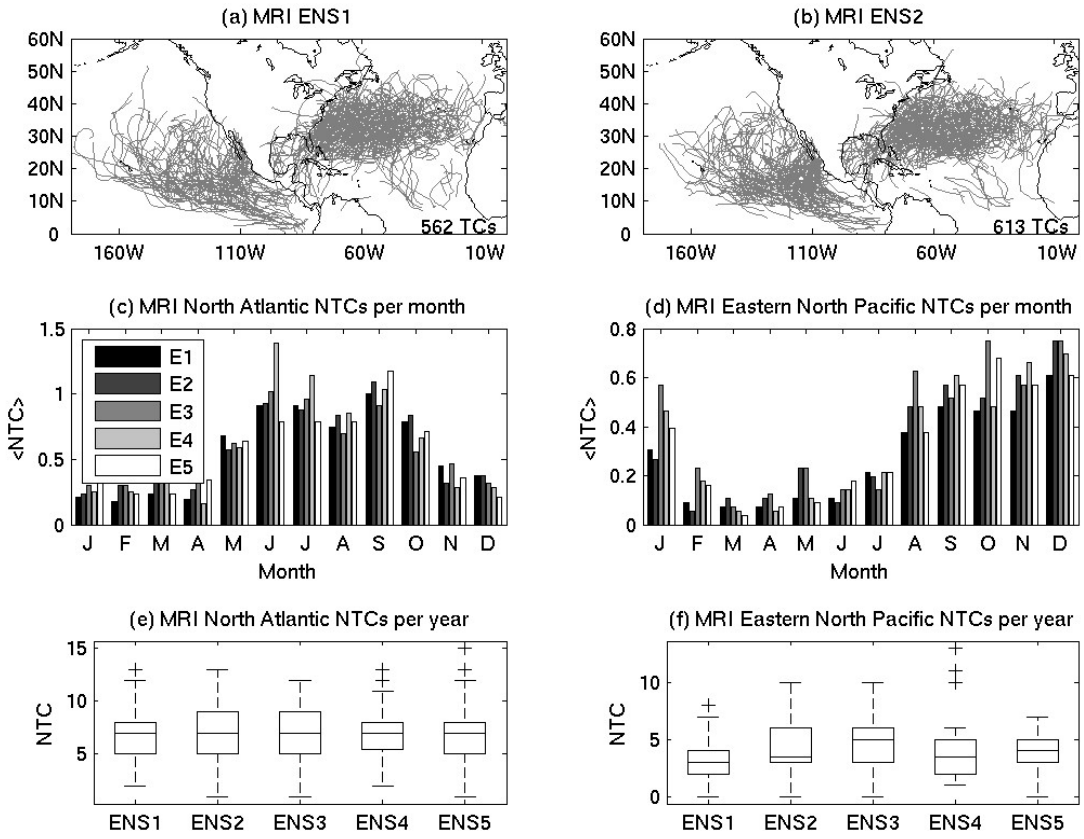

FIG. 9: Tracks of the MRI model TCs for ensemble members (ENS) 1 (a) and 2 (b). Mean MRI NTC per month in North Atlantic (c) and eastern North Pacific (d) for 3 ensemble members (E1, E2, E3, E4, and E5) in the period 1950-2005. Distributions of MRI NTC per year for 5 ensemble members in the North Atlantic (e) and eastern North Pacific (f) in the period 1950-2005. The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

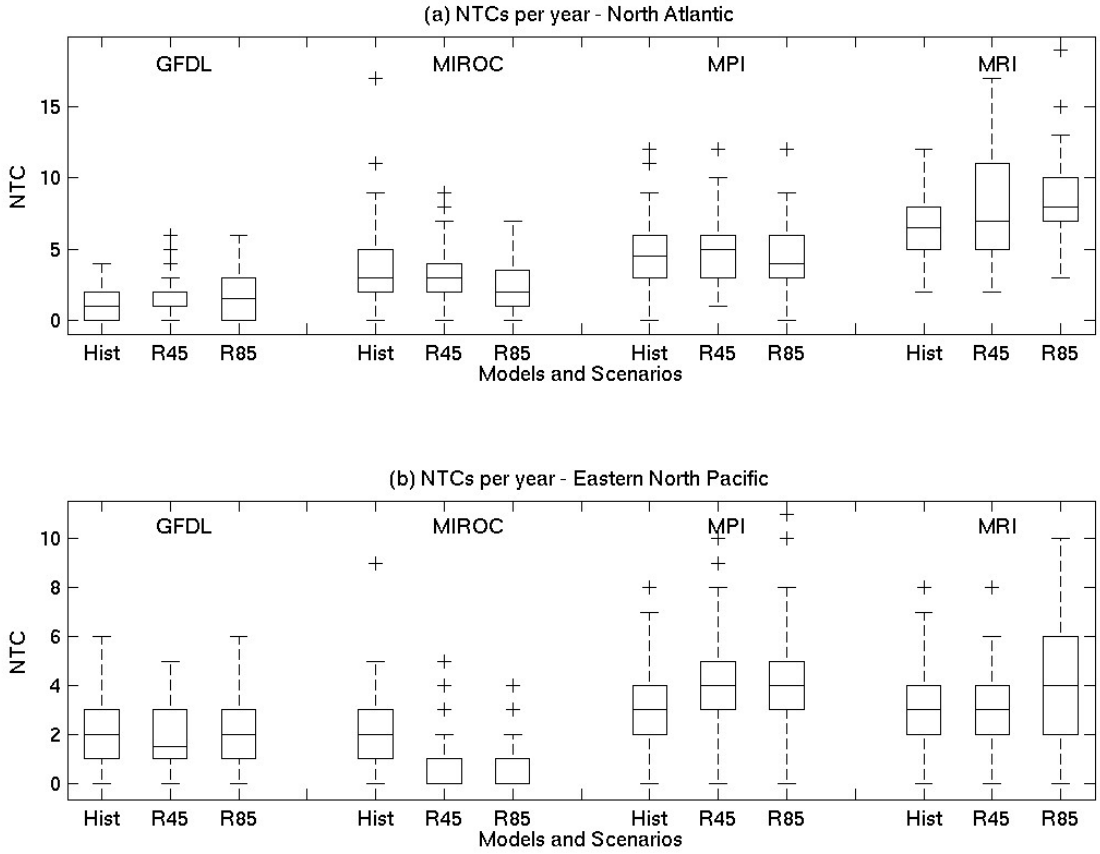

FIG. 10: NTC per year in the models in North Atlantic (top panel) and eastern North Pacific (bottom panel) for the historical (Hist; 1951-2000) and RCP4.5 (R45) and RCP8.5 (R85) future scenarios (2051-2100). The box denotes the range of the 25th to 75th percentiles of the distributions, with the median marked by the line inside the box and the values outside of the middle quartile being marked by whiskers and crosses.

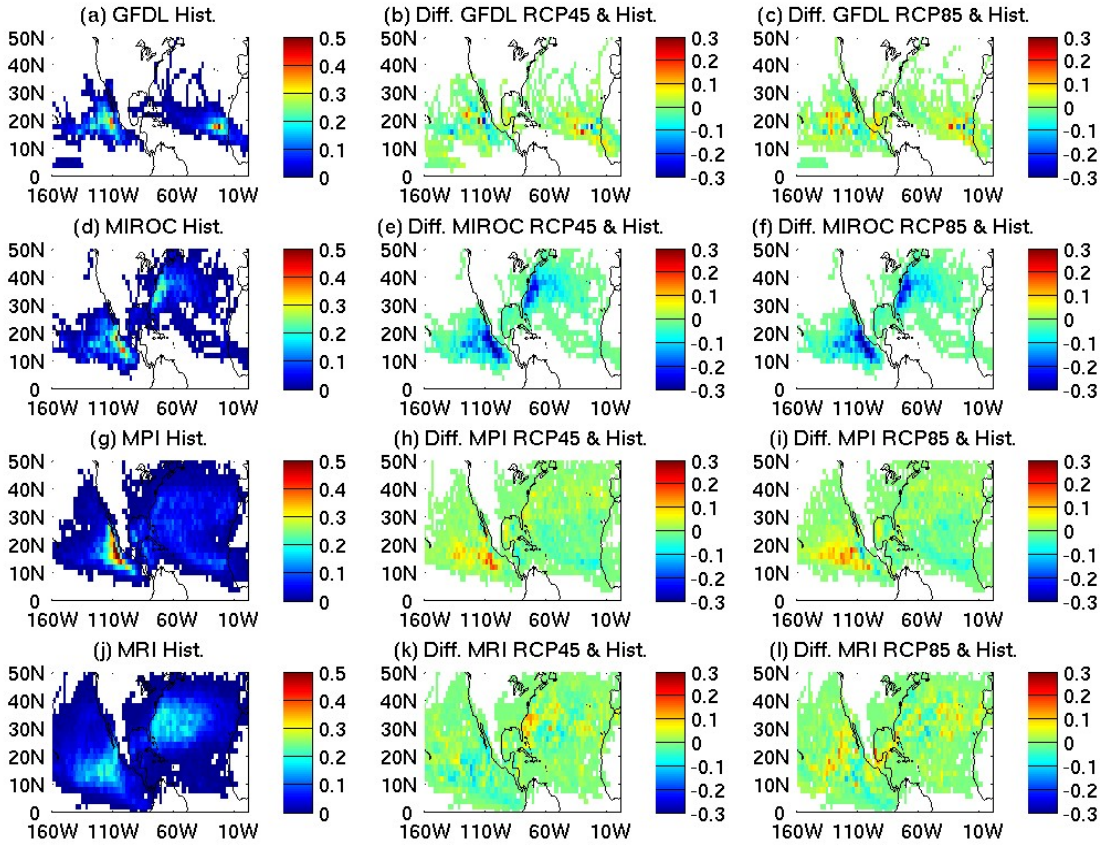

FIG. 11: Model climatological track density (mean counts of tracks passing per grid point) in historical runs (left panels). Difference of the model climatological track density for the future scenarios and historical run: RCP4.5 (middle panels) and RCP8.5 (right panels). In the cases that more than one ensemble member is available, all ensembles were used to calculate the climatological track density.

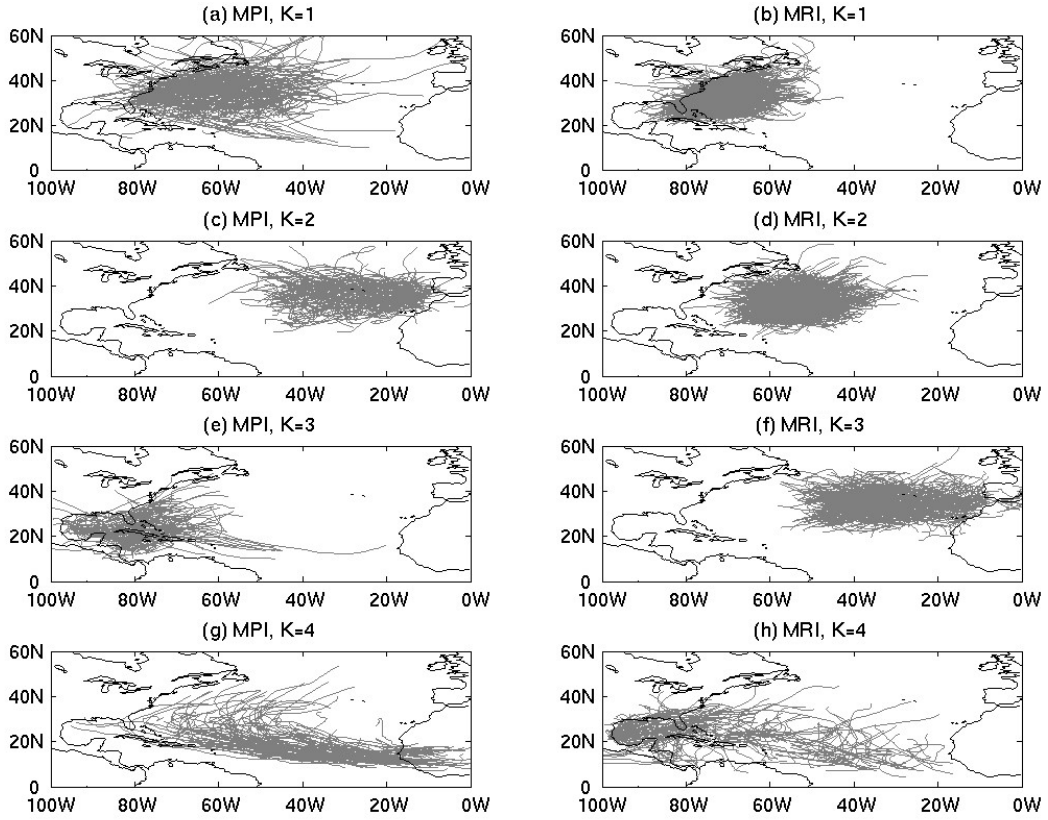

FIG. 12: North Atlantic TCs tracks by cluster (K) for the MPI (left panels) and MRI (right panels) models. Tracks for one ensemble member of the historical, RCP4.5 and RCP8.5 runs are shown together for each cluster and model.

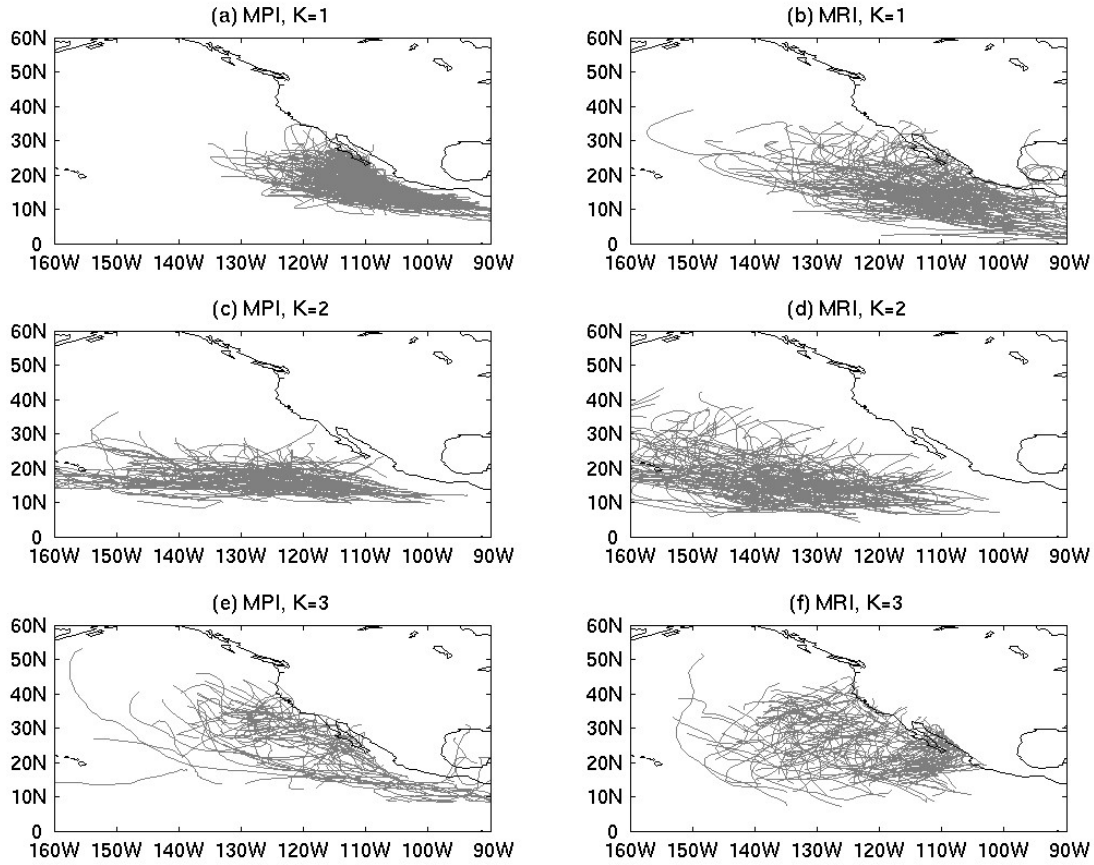

FIG. 13: Eastern North Pacific TCs tracks by cluster for the MPI (left panels) and MRI (right panels) models. Tracks for one ensemble member of the historical, RCP4.5 and RCP8.5 runs are shown together for each cluster and model.

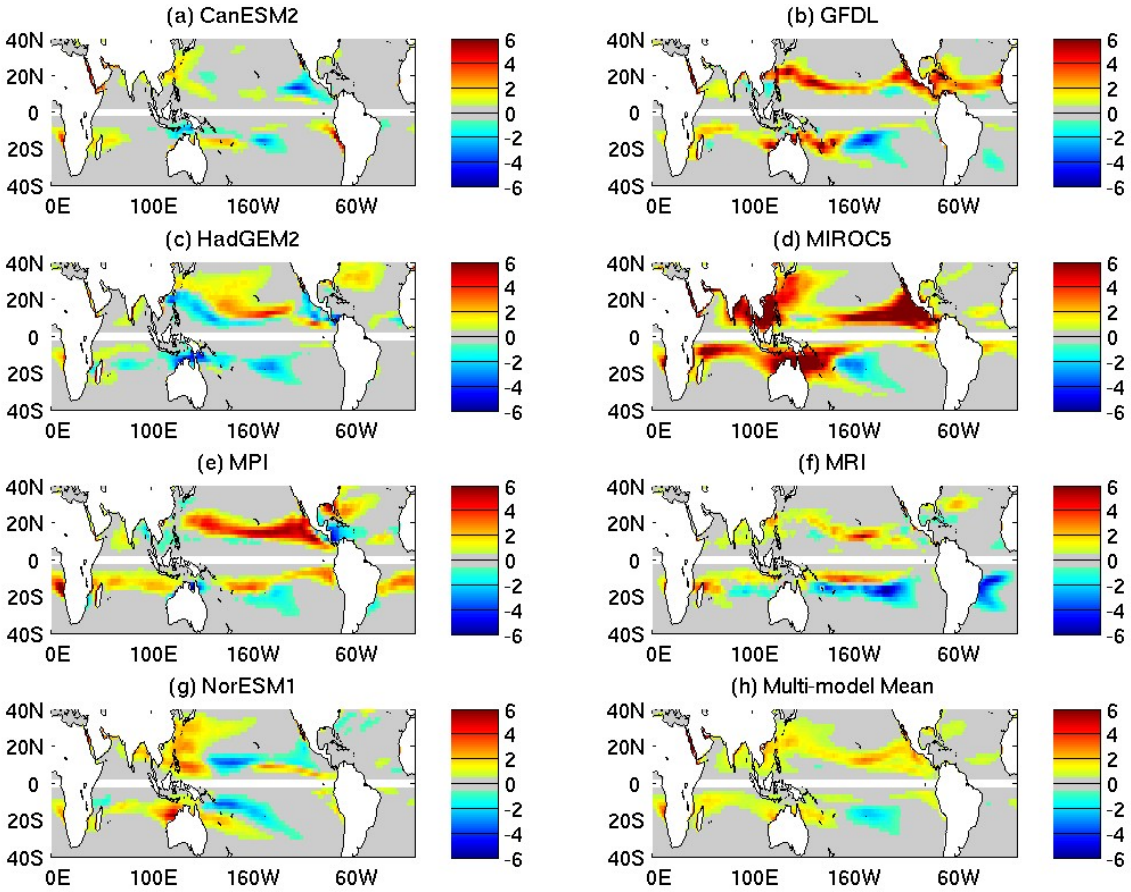

FIG. 14: Difference of model genesis potential index (GPI) climatology between the RCP8.5 future scenario (2071-2100) and historical (1971-2000) run. The multi-model mean difference is shown in the right bottom panel. In all panels, the difference in ASO (JFM) is shown in the northern (southern) hemisphere. All ensemble members available are used to calculate the GPI climatology per model and scenario.

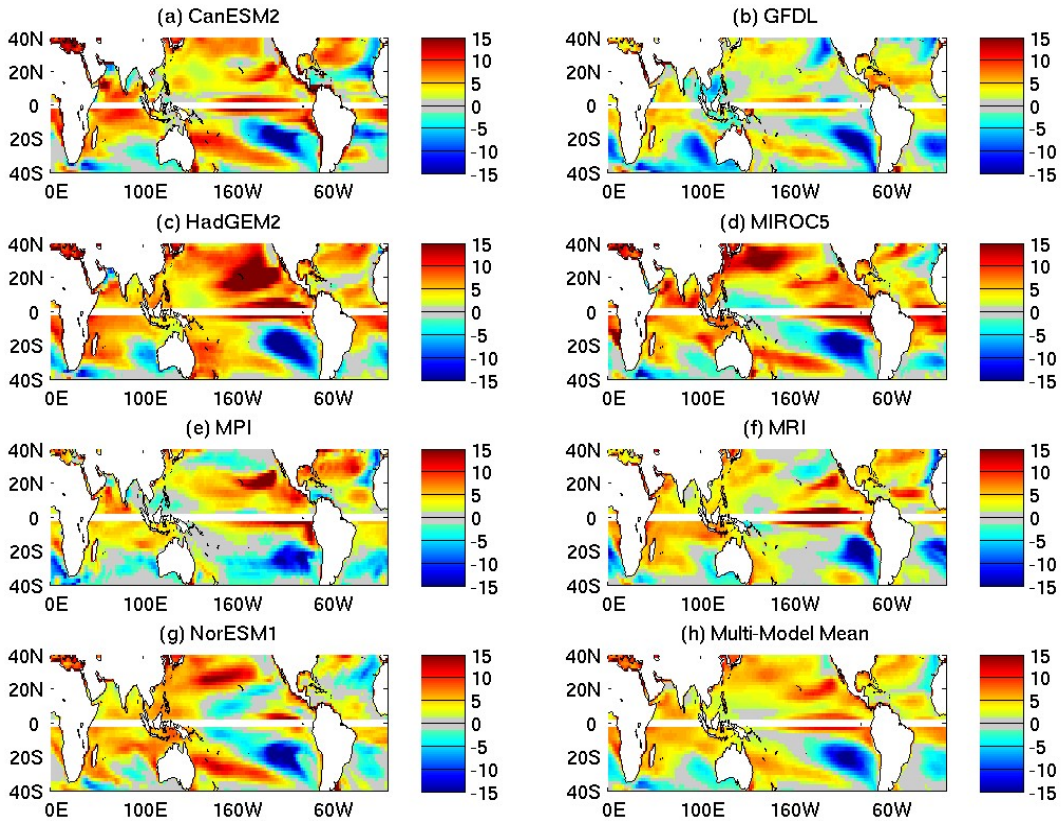

FIG. 15: Difference of model potential intensity (PI) climatology between the RCP8.5 future scenario (2071-2100) and historical (1971-2000) run. The multi-model mean difference is shown in the right bottom panel. In all panels, the difference in ASO (JFM) is shown in the northern (southern) hemisphere. All ensemble members available are used to calculate the PI climatology per model and scenario.

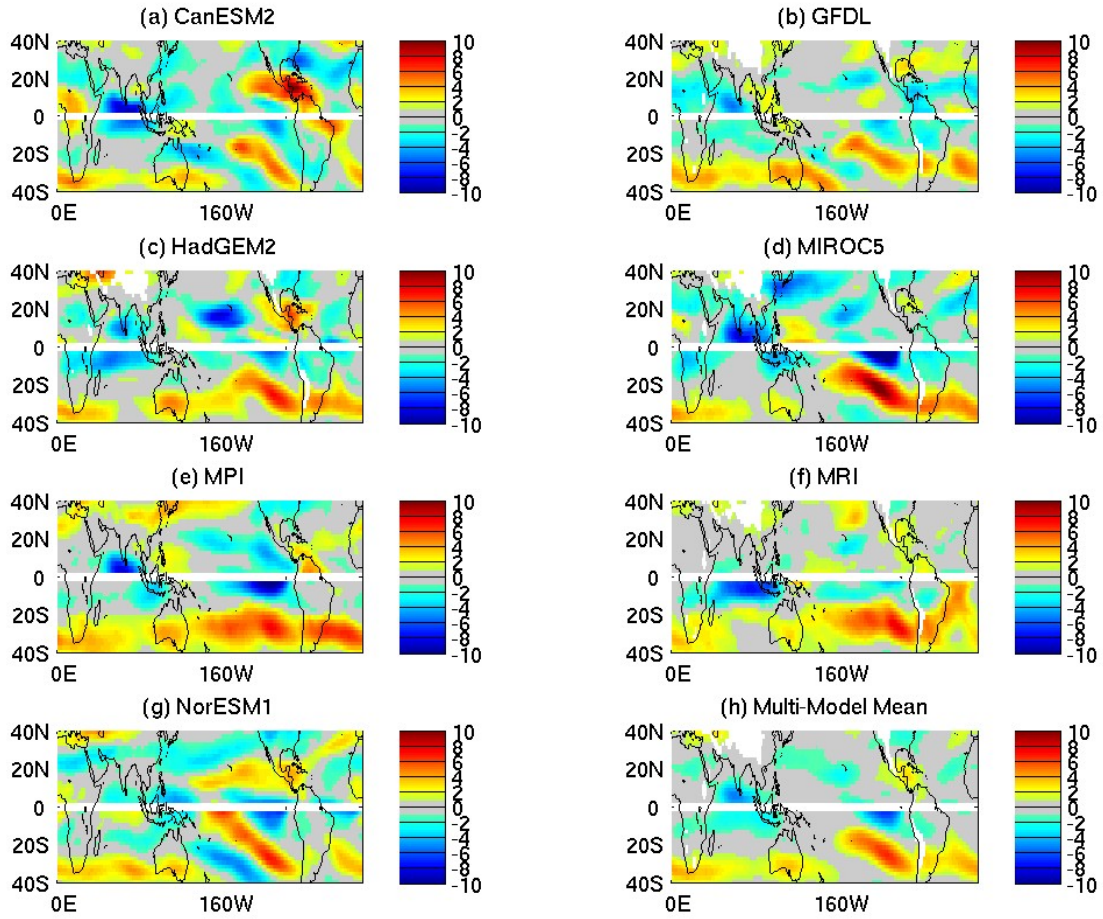

FIG. 16: Difference of climatological model vertical wind shear between the RCP8.5 future scenario (2071-2100) and historical (1971-2000) run. The multi-model mean difference is shown in the right bottom panel. In all panels, the difference in ASO (JFM) is shown in the northern (southern) hemisphere. All ensemble members available are used to calculate the magnitude of the vertical wind shear climatology per model and scenario. The vertical winds shear magnitude is calculated as the magnitude of the difference between the 200hPa and 850hPa winds.

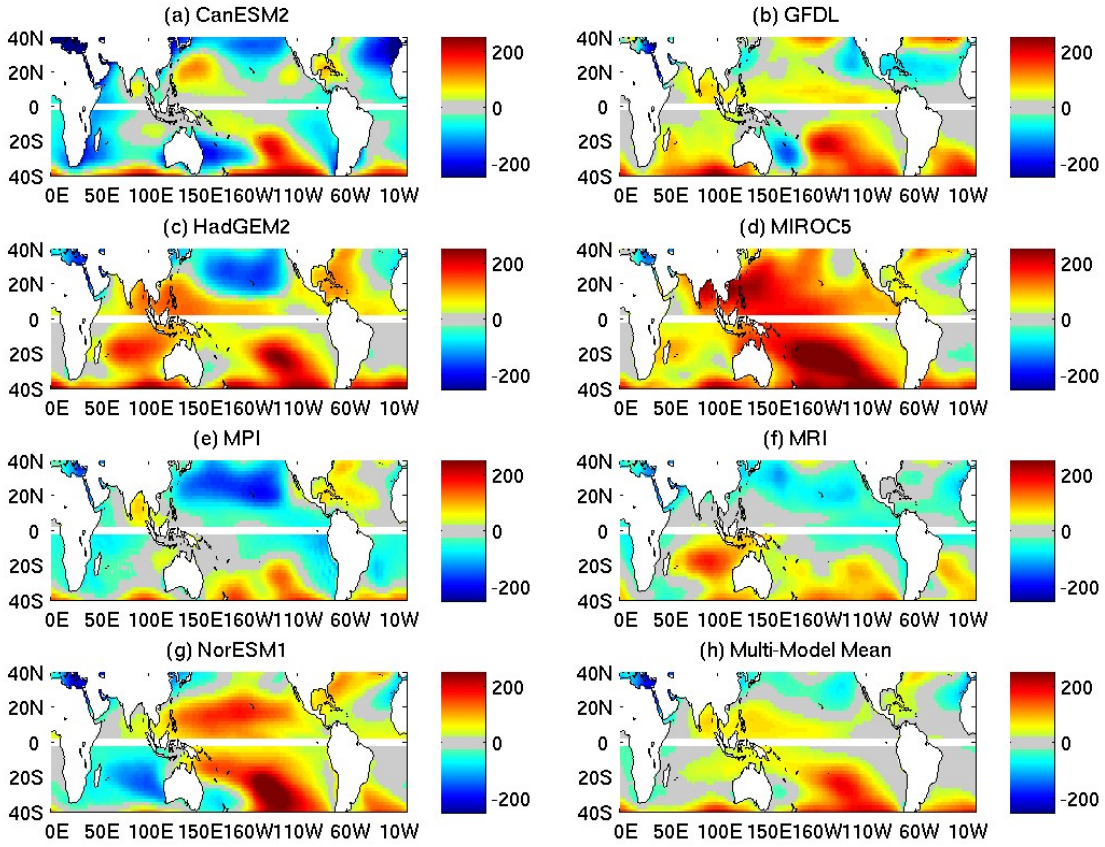

FIG. 17: Difference of climatological sea level pressure between the RCP8.5 future scenario (2071-2100) and historical (1971-2000) run. The multi-model mean difference is shown in the right bottom panel. In all panels, the difference in ASO (JFM) is shown in the northern (southern) hemisphere. All ensemble members available are used to calculate the sea level pressure climatology per model and scenario.